GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

January 6, 2004, 01:15:17; Search time 2583 Seconds Run on:

(without alignments)

9541.130 Million cell updates/sec

Delakora i

US-10-089-872-1 Title:

Perfect score: 1014

1 atgaaaaaaatgcctttgtt.....tgaagaaaacggccccttga 1014 Sequence:

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

22781392 seqs, 12152238056 residues Searched:

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Fort processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

EMENDESIS I

EST:*

1: em_estba:*

2: em_esthum:*

3: em estin:*

4: em estmu:*

5: em_estov:*

v: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:* 12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_esttun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:* 19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*
26: em_gss_phg:*

27: em gss_vrl:*

28: gb_gss1:* 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

						DOMMET	30
			용				
Result	t		Query				Description
No.		Score	Match	Length	DB	ID	Description
		-	 -				AK076867 Mus muscu
	î.	260.4	84.9	1552	11	AK076867	AK030474 Mus muscu
2	2	860.4	84.9	2245	11	AK030474	AK053642 Mus muscu
:	3	360.4	84.9	3039	11	AK053642	AK075758 Mus muscu
4	4	358.8	84.7	1377	11	AK076758	AK013205 Mus muscu
	5	8.445	83.3	1449	11	AK013205	AU125107 AU125107.
(6	770.6	76.0	822	9	AU125107	BX393735 BX393735
	7	750.8	74.0	1201	13	BX393735	AK005323 Mus muscu
	3	709.4	70.0	1379	11	AK005323	BG218735 RST38476
	9	571.8	66.3	784	10	BG218735	AK013161 Mus muscu
1	0	622.8	61.4	1281	11	AK013161	BU116522 603139786
7	<u>:</u>	614	60.6	951	13	BU116522	BQ669953 AGENCOURT
1	2	94.2	58.6	982	13	BQ669953	BU518807 AGENCOURT
-1		535.2	57.7	934	13	BU518807	CD354831 UT-M-GMO-
3.	4	79.2	57.1		14	CD354831	Bx483012 Homo sapi
1.	5	578	57.0	742	2	HSM073180	CA973078 AGENCOURT
٦.	.6	565	55.7		1.4	CA973078	BU514920 AGENCOURT
0	7	353.4	54.6	958	1.3	BU514920	AA278473 2881h12.T
	.3	550.27	54.3			AA278473	CA982682 AGENCOURT
1	.9	542.5	53.5			CA982682	CD103801 AGENCOURT
1	0	536.2	52.9			CD103801	BF159587 601769084
33	1	533.2	52.6			BF159587	AW242839 xn26f05.x
c. 🗟	2	521.4	51.4			AW242839	BU914186 AGENCOURT
2	:3	537.8	51.1			BU914186	BQ899617 AGENCOURT
3	.4	500.4	49.3			BQ899617	CB559060 AGENCOURT
2	25	499.6	49.3			CB559060	BU367571 603787711
2	26	491.8	48.5				BQ940647 AGENCOURT
2	27	486.6	48.0				CB200465 AGENCOURT
2	2.8	473.6	46.7				BQ021348 UI-H-DH1-
c 2	.9	461	45.5				CA983082 AGENCOURT
3	0 8	460.8	45.4				BI499153 ie27c05.y
3	31	157.6	45.1				BU905264 AGENCOURT
3	52	.452	44.6				BI655225 603284039
-	3.3	451.4	44.5			- BI655225	BQ880200 AGENCOURT
ء در	34	451.2	44.5				BU152376 AGENCOURT
	35	445.2	43.9				AI645170 mt82a02.y
3	36	445	43.9			AI645170	
-	37	443.8	43.8			AL961440	AL961440 AL961440 CB564338 AGENCOURT
c c	38	442.8	43.	7 853			BI828691 603074707
	39	441.6	43.6	695	7 12		CB235478 AGENCOURT
2	4. C	435.8	43.0	1046	5 14		
2	4, 1	435.2	42.5			AA669484	AA669484 af74g07.r BU631151 UI-H-FE1-
c '	4.2	434.2	42.8				BU129770 603117261
	43	433	42.				CONTRACTOR AND THE MENT
	44	430	42.				CT OF OR O. ACTINGOLDE
	45	428	42.	2 98	5 14	CA970822	CA9/0822 AGENCOURT

ALIGNMENTS

```
RESULT 1
AK076867
                                                                    HTC 07-DEC-2002
                                      1552 bp
                                                  mRNA
                                                           linear
LCCUS
             AK076867
            Mus musculus adult male testis cDNA, RIKEN full-length enriched
DEFINITION
             library, clone:4930520C08 product:MO25-LIKE PROTEIN homolog [Homo
             sapiens], full insert sequence.
            AK076867
ACCESSION
            AK076867.1 GI:26345723
VERSION
KEYWORDS
            HTC; CAP trapper.
            Mus musculus (house mouse)
SOURCE
            Mus musculus
  ORGANISM
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
             Carninci, P. and Hayashizaki, Y.
  AUTHORS
            High-efficiency full-length cDNA cloning
  CITLE
            Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
             99279253
  MEDLINE
             10349636
   ₽UBMED
             2
REFERENCE
             Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  MUTHORS
             Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
            Wormalization and subtraction of cap-trapper-selected cDNAst to
  TENER.
             prepare full-length cDNA libraries for rapid discovery of new genes.
                                                                          , i, . 2
             Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
                                                                          146
             20499374
  MEDLINE
 DUBMED
            11042159
                                                                       C_{i}\subseteq C_{i}
REFERENCE
             Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  RUTHORS
             Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Iton, M.,
             Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
             Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
             Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
             Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
             Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
  TITLE
             sequencing pipeline with 384 multicapillary sequencer
             Genome Res. 10 (11), 1757-1771 (2000)
  JOURNAL
  MEDLINE
             20530913
             11076861
  PUBMED
REFERENCE
             Kawai, J., Shinaqawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
AUTHORS
             Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
             Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
             Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
             Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
             Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
             Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
             Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
             Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
             Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
```

Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,

```
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
            Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
            Marchionni, I.., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
            Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
            Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
            Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
            Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
            and Hayashizaki,Y.
            Functional annotation of a full-length mouse cDNA collection
  TITLE
            Nature 409 (6821), 685-690 (2001)
  JOURNAL
            21085660
  MEDLINE
            11217851
   PUBMED
REFERENCE
            The FANTOM Consortium and the RIKEN Genome Exploration Research
  AUTHORS
            Group Phase I & II Team.
            Analysis of the mouse transcriptome based on functional annotation
  TITLE
            of 60,770 full-length cDNAs
            Nature 420, 563-573 (2002)
  JOURNAL
                (bases 1 to 1552)
REFERENCE
            Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
  AUTHORS
            Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
            Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
            Miraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
            Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
            Yoya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
            Nemura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
            Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
            Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
            Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
             Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,
             Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
             Direct Submission
  TITLE
             Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
  JOURNA"...
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             WRL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             %ax:81-45-503-9216)
             cDNA library was prepared and sequenced in Mouse Genome
 COMMENT
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
             Please visit our web site for further details.
             URL:http://genome.gsc.riken.go.jp/
             URL:http://fantom.gsc.riken.go.jp/.
                       Location/Qualifiers
 FEATURES
                       1. .1552
      source
                       /organism="Mus musculus"
                       /mol_type="mRNA"
                       /strain="C57BL/6J"
                       /db_xref="FANTOM_DB:4930520C03"
                       /db_xref="MGI:1894876"
                       /db xref="taxon:10090"
                       /clone="4930520C03"
```

/sex="male"

```
/tissue_type="testis"
               /clone_lib="RIKEN full-length enriched mouse cDNA library"
               /dev stage="adult"
               316. .1320
   CDS
               /note="unnamed protein product; MO25-LIKE PROTEIN homolog
               [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
               98.2%ID, 100%length, match=1002)
               putative"
               /codon start=1
               /protein_id="BAC36513.1"
               /db xref="GI:26345724"
               /db xref="MGI:1914081"
               translation="MPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQ/
               AMKEILCGTNDKEPPTEAVAQLAQELYSSGLÜVTLIADLQLIDFEGKKDVTQIFNNIL
               RRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIILFSN
               QFRDFFKYVELSTFDTASDAFATFKDLLTRHKVLVADFLEQNYDTTFEDYEKLLQSEN
               YVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVFKV
               FVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKKAAP
               1539. .1544
   polvA signal
               /note="putative"
               1552
   polyA_site
               /note="putative"
                        341 g
                               401 t
           490 a
                  320 €
BASE COUNT
ORIGIN
                                         Length 1552;
                         Score 860.4; DB 11;
                   84.9%;
 Query Match
                        Pred. No. 5.8e-176;
 Best Local Similarity
                   90.5%;
                                        Indels
                                                  'Gaps. @
                           Mismatches
                                    96:
                        0 ;
 Matches 918; Conservative
        1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
QΥ
          307 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAAATC 366
D:
        Qγ
                   Db.
       QΥ
          Db
       1.81 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
QУ
          437 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 546
ac.
       241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
Qy
          547 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 606
Db-
       F01 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
ČΛ
          607 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 666
Db
       361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
Qу
           667 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 726
Db
```

Qy 421	CGTTGTGGGATTATGCTGAGAAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 4	80
Dò 727	CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 7	86.
Qy 431	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 5	4 O·
Db 787	TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 8	46
Qy 541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 6	00
Db 847	TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 9	06
Qy 601	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 6	60
. Db 907	TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 9	66
Q <u>v</u> 661	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 7	20.
Db 967	AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 1	026
Qy 721	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 7	80
Db27	AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC	086
701	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 8	40:
Dr. 1937	O FGCTTCGAGACAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1	146
Gy 841	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 9	0.0
Db 1147	GTGGCCAGCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1	2.06
Qy 901	CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 9	60
Db 1207	CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1	266
Оу 961	GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014	•
Db 1267	GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1320	a tr
manarana sa ka		
RESULU C AK030474		
LOCUS	AK030474 2245 bp mRNA linear HTC 05-DEC-2	
	Mus musculus adult male pituitary gland cDNA, RIKEN full-length a enriched library, clone:5330416K15 product:MO25-LIKE PROTEIN	
	nomolog [Homo sapiens], full insert sequence.	
	AK030474 AK030474.1 GI:26326468	
	HTC; CAP trapper.	
	Mus musculus (house mouse)	
E	dus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi	
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mu	s.
REFERENCE 1 AUTHORS (l Carninci,P. and Hayashizaki,Y.	

```
High-efficiency full-length cDNA cloning
  TITLE
            Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
            99279253
  MEDLINE
            10349636
   PUBMED
            2
REFERENCE
            Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  AUTHORS
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
  TITLE
            prepare full-length cDNA libraries for rapid discovery of new genes
            Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
            20499374
  MEDLINE
            11042159
   PUBMED
REFERENCE
            Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  AUTHORS
             Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
             Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A..
             Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
             Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
             Yoneda, Y. Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
             Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
  TITLE
             sequencing pipeline with 384 multicapillary sequencer
             Genome Res. 10 (11), 1757-1771 (2000)
  COURNAL
             20530913
  MEDLINE
   DUBMED
             11076861
PREBENCE
             Mawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Sh
  MILCRS
             Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., A.
             Aizawa, K., Izawa, M., Wishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
             Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
             Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
             Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
             Wuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
             Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
             Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
             Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
             Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
             Eletcher, C., Fujita, M.; Gariboldi, M., Gustincich, S., Hill, D.,
             Mofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
             Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
             Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
             Sato, K., Schonbach, C., Seya. T., Shibata, Y., Storch, K.F., Suzuki, H.,
             Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
             Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
              and Hayashizaki, Y.
             Functional annotation of a full-length mouse cDNA collection
   THTLE
             Nature 409 (6821), 685-690 (2001)
   JOURNAL
   MEDLINE
              21085660
              11217851
    PURMED
 REFERENCE
              The FANTOM Consortium and the RIKEN Genome Exploration Research-
   AUTHORS
              Group Phase I & II Team.
              Analysis of the mouse transcriptome based on functional annotation
   TITLE
              of 60,770 full-length cDNAs
              Nature 420, 563-573 (2002)
   JOURNAL
                 (bases 1 to 2245)
 REFERENCE
              Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
   AUTHORS
```

```
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
      Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
      Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
      Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
      Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
      Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
      Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
      Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
      Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
      Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
      Muramatsu, M. and Hayashizaki, Y.
      Direct Submission
      Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
       Physical and Chemical Research (RIKEN), Laboratory for Genome
       Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
       RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
       Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
       URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
       Fax:81-45-503-9216)
       cDNA library was prepared and sequenced in Mouse Genome
       Encyclopedia Project of Genome Exploration Research Group in Riken
       Genomic Sciences Center and Genome Science Laboratory in RIKEN.
       Division of Experimental Animal Research in Riken contributed to
       prepare mouse tissues.
       Please visit our web site for further details.
       URL:http://genome.gsc.riken.go.jp/
       IRL:http://fantom.gsc.riken.go.jp/.
                Location/Qualifiers
                1. .2245
ವರಿಗಳ ನಿಜ
                /organism="Mus musculus"
                /mol type-"mRNA"
                /strain="C57BL/6J"
                /db xref="FANTOM_DB:5330416K15"
                /db xref="taxon:10090"
                /clone="5330416K15"
                /sex="male"
                /tissue_type="pituitary gland"
                /clone_lib="RIKEN full-length enriched mouse cDNA library"
                /dev_stage="adult"
                300. .1313
                /note="unnamed protein product; MO25-LIKE PROTEIN homolog
                 [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                98.2%ID, 100%length, match=1002)
                putative"
                /codon start=1
                /protein_id="BAC26978.1"
                /db xref="GI:26326469"
                translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK/
                SLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
                NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
                FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
                SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
                FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKK
                AAP"
```

581 t

480 c · 517 g

TITLE.

COMMENT

RES

CDS

BASE COUNT

ORIGIN

667 a

JOURNAS

	Query Ma Best Loc Matches	al S	imilarity; Conserva	90.5%;	Pred.	No. 6.26	e-176;	Length Indels	2245; 0;	Gaps	0.;
Q: D:			ATGAAAAAAAT ATGAAAAAAAT		11111			1	1111		
Q: D:	-		CTGAAAGACAA CTGAAAGACAA	111111	++++++		111111			11111	
D D	У	121	GAAGTGTCTAA GAGGTGTCAAA	ATCACTGC	AAGCAAT	rgaaagaa 	ATTCTGT(GTGGTACA/	AACGAG	AAAGAA	180
Q	У	181	CCCCCAACAGA	AGCAGTGG	CTCAGC	ragcacaa 	GAACTCT.	ACAGCAGT(GCCTG	CTAGTG	240 .
Ç	Y	241	ACACTGATAGO	TGACCTGC	CAGCTGA	ragacttt 	: GAGGGAA 	AAAAAGAT(GTGACC	CAGATA	300
Q	y Db	303	TTTAACAACAT	CTTGAGAA	AGACAGA'	TAGGCACT	CGGAGTC	CTACTGTG	GAGTAŢ	'ATTAGT	360
	ist Ser		GCTCATCCTCA				11 4111	111111			
:	TY DS		CGTTGTGGGAT	11111	11111 1	1 {		11111	11111		
	Ω y 	481	TTTTCTAATCA	AATTCAGAG	GATTTCT	TTAAGTAC	GTGGAGT	TTGTCAACA	TTTGA'	TATTGCT	54.0
	Qy Ob		TCAGATGCCT		11 1 1				11111		
	DP GA		TTCTTAGAAC		11111				11 11		
	Ο Σ Σ		AATTATGTTA	1 11111			1111	i		11111	
,	Qy Db		AACTTTGCCA	1 1111		[11111	
	Qу		CTCCTTCGGG								

```
1030 CTGCTTCGAGACAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1139
Db
         841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
Qу
             1140 GTGGCCAGCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1199
Db
       901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
QУ
             1200 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGACGACGACGAGCAGCTTTGCTGAC 1259
Db
         961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Qy.
             1269 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1313
Db
RESULT 3
AK053642
                                                            HTC 05-DEC-2002
                                            mRNA
                                                   linear
                                  3039 bp.
           AK053642
LOCUS
          Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
DEFINITION
           library, clone:E130116021 product:M025-LIKE PROTEIN homolog [Homo
           sapiens], full insert sequence.
           AK053642
ACCESSION
           AK053642.1 GI:26343600
VERSION
           HTC; CAP trapper.
KEYWORDS
           Mus musculus (house mouse)
COURCE
  CRGANIES Mus musculus
           Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE:
           Carninci, P. and Hayashizaki, Y.
  AUTHORS
           High-efficiency full-length cDNA cloning
  CHRLE
           Meth. Enzymol. 303, 19-44 (1999)
                                                                 1000
  TRNA L
                                                                V 60
           99279253
  MEDLINE
           40349636
   SJEMED .
REFERENCE
           Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  AUTHORS
           Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
           Normalization and subtraction of cap-trapper-selected cDNAs to
  TITLE.
           prepare full-length cDNA libraries for rapid discovery of new genes
           Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
           20499574
  MEDLINE
           11042159
   PUBMED
REFERENCE
           Shibata, K., Iton, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  AUTHORS
           Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
           Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
           Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
           Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
           Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
           Okazaki.Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
           RIKEN integrated sequence analysis (RISA) system--384-format
  TITLE
            sequencing pipeline with 384 multicapillary sequencer
           Genome Res. 10 (11), 1757-1771 (2000)
  JOURNAL
           20530913
  MEDITINE
           11076861
   PUBMED
REFERENCE
```

```
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 AUTHORS
           Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
           Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
           Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
           Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
           Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
           Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
           Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
           Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
            Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
            Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
            Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
            Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
            Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
            Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
            Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
            and Hayashizaki, Y.
            Functional annotation of a full-length mouse cDNA collection
  TITLE
            Nature 409 (6821), 685-690 (2001)
  JOURNAL
            21085660
 MEDLINE
            11217851
  PUBMED
REFERENCE
            The FANTOM Consortium and the RIKEN Genome Exploration Research
  AUTHORS
            Group Phase I & II Team.
            Analysis of the mouse transcriptome based on functional annotation
  CITLE
            of 50,770 full-length cDNAs ...
            Mature 420, 563-573, (2002)
  JOURNAL
                (bases 1 to 3039)
            Adachi, J., Aizawa, K., 'Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
REFERENCE
  AUTHORS
            Fukuda.S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
            Mayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
            Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
            Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
            Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
            Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
            Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
            Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
             Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
             Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
             Muramatsu, M. and Hayashizaki, Y.
             Direct Submission
  TITLE
             Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
  JOURNAL
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            PRIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             Fax:81-45-503-9216)
             cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
COMMENT
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
```

prepare mouse tissues.

Please visit our web site for further details.

URL:http://genome.gsc.riken.go.jp/

```
URL: http://fantom.gsc.riken.go.jp/.
               Location/Qualifiers
FEATURES
                1. .3039
   source
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6J"
                /db_xref="FANTOM_DB:E130116021"
                /db xref="taxon:10090"
                /clone="E130116021"
                /tissue_type="eyeball"
                /clone lib="RIKEN full-length enriched mouse cDNA library"
                /dev_stage="0 day neonate"
                287. .1300
   CDS
                /note="unnamed protein product; MO25-LIKE PROTEIN homolog
                [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                98.2%ID, 100%length, match=1002)
                putative"
                /codon start=1
                /protein_id="BAC35457.1"
                /db_xref="GI:26343601"
                /translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
                SLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
                NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
                FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
                SENYVTKROSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
                FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKK
                AAP"
                3022. .3027
    golyA_signal
                /note="putative"
                3039
    folyA_site
                /noce="putative"
                   645 C
                          676 P
            882 a
BASE COUNT
ORIGIN
                          Score 860.4; DB 11: Length 3039;
                    84.98;
  Quary Match
                          Pred. No. 6.6e-176;
                    90.5%;
  Boot Local Similarity
                                          Indels
                                                     Gaps
                         0; Mismatches
                                      96;
  Matches 918; Conservative
         1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60.
ŨΨ
           237 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 346-
Db
         Qy
                    File
        Qу
           Db
        181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Çy
           467 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 526
Db
        241 ACACTGATAGCTGACCTGCACCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300.
 Qу
           527 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 586
 Db
```

_ -	3.0.1	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	360
Ďλ		TTTAACAACATCTTGAGAAGACTGTTTTAACAACATCTTGAGAAGACTGTTGTTGAGAACATCTTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT	
)b		TO THE TAX TO THE TRANSPORT OF THE TAX A GGA TA TGA AGCCCCACAGATTGCCTTA	
ДУ		GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGTTTTTTTT	
ďù		CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	
Qy Db		CGTTGTGGGATTATGCTGAGAGAATGTATTGGACHTGATATGTATTGTAGAGATGTATTGGACATGTATTGTAGAGAGAG	
Qу		THE ATTENDANT OF THE CASE ATTENDED TO THE ATTE	
Db	767	TTTTCTAATCAGTTCAGAGATTTCTTTAAGTTGGTGTGTGT	826 🤄
Qv	541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	600
ි. වසි :	827	TCAGATGCCTTTGCTACTTTCAAGGATTTACTATCGTAGCTACTTTGCTACTTTTAAGGATTTTAACCAGACATAAAGTATTGGTAGCAGACATAAAGTATTGGTAGCAGACAGA	886
ДÄ		TO SEA OF CHARTETTE A CACTATGAGAAATTGCTTCAGTCTGAG	
Db	8.87	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAATTACGACACTATTTTTGAAGACTATGAGAAAATTATGACACTATTTTTTGAAGACTATGAGAAAACTGCTGCAATCTGAG	946 :
99		THE COLOR OF CHARGE CONTROL AND CONTROL AN	
775)		ARTTATGTTACTAAGAGACAGTCTTTAAAGCTGCTTTGGAGCTGATCCTGGACCGCCAC	
Qy ·		TO TO TO TO TO TO TAKE A CARACTATATATA TO A COLOR A CO	780%
794 794		AACTTTGCCATCATGACAAAGTATATCAGCAAGCGGATATATCAGCAAGCGGATGATGATGACGAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAACCAAGCCAGAGAACCTGAAACTGAAACTGATGAAACTGATGAAACTGAAACTGAAACTGAAACTGAAACTGAAAACTGAAACTGAAACTGAAACTGAAACTGAAACTGAAACTGAAACTGAAACTGAAACTGAAACTGAAACTGAAACTGAAACTGAAACAACAAACA	
√y.		THE REPORT OF THE PARTY OF THE	840
no no	3.065	CTCCTTCGGGATAAAAGTCCCAACATGTTTTTTTTTTTT	1126
Qy		THE STATE OF THE S	4 90:0 ±
 تb		GTGGCCAGTTCTCACAAAACACAGCTATTGTGGAGATTGTTGTGGAGATTGTGTGTG	l .
		CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAG	
Ωγ Ελ·			
·Qyi	. "	1 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014	5,4
, ∖ex ∵Db			
DD		•	

AK076758

LCCUS AK076758

LCCUS AK076758

AK076758

1377 bp mRNA linear HTC 07-DEC-2002

Mus musculus adult male testis cDNA, RIKEN full-length enriched

Library, clone:4930433N18 product:MO25-LIKE PROTEIN homolog [Homo sapiens], full insert sequence.

ACCESSION AK076758

AK076758.1 GI:26345637 VERSION HTC; CAP trapper. KEYWORDS Mus musculus (house mouse) SOURCE Mus musculus ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. REFERENCE Carninci, P. and Hayashizaki, Y. AUTHORS High-efficiency full-length cDNA cloning TITLE Meth. Enzymol. 303, 19-44 (1999) JOURNAL 99279253 MEDLINE 10349636 DIJBMED REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., AUTHORS Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to TITLE prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) JOURNAL MEDLINE 20499374 PUBMED 11042159 REFERENCE Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., AUTHORS Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format EILTE sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) COURNAL 20530913 DLINE PUBMED 11076861 MEMENCE Hawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., ACTHORS. Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Pleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection TITLE Nature 409 (6821), 685-690 (2001) JOUENAL. 21085660 MEDLINE

11217851

PUBMED

```
The FANTOM Consortium and the RIKEN Genome Exploration Research
REFERENCE
  AUTHORS
            Group Phase I & II Team.
            Analysis of the mouse transcriptome based on functional annotation
  TITLE
            of 60,770 full-length cDNAs
            Nature 420, 563-573 (2002)
  JOURNAL
                (bases 1 to 1377)
REFERENCE
            Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
            Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
  AUTHORS
            Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
            Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
            Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
            Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
            Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
            Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
            Shibata K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
            Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
            Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,
            Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
             Direct Submission
             Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
  TITLE
  JOURNAL
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             TRL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
                                                                                   <u>F.</u>
             max:81-45-503-9216)
              ONA library was prepared and sequenced in Mouse Genome
                                                                                   CL
             Encyclopedia Project of Genome Exploration Research Group in Riken e
 COMMENT
             denomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
             Please visit our web site for further details.
             URL:http://genome.gsc.riken.go.jp/
             URL:http://fantom.gsc.riken.go.jp/.
                       Location/Qualifiers
 TEATURES
                       1. .1377
      source
                       /organism="Mus musculus"
                       /mol_type="mRNA"
                       /strain="C57BL/6J"
                       /db_xref="FANTOM_DB:4930433N18"
                       /db_xref="MGI:1894867"
                       /db xref="taxon:10090"
                       /clone="4930433N18"
                       /sex="male"
                       /tissue_type="testis"
                       /clone_lib="kIKEN full-length enriched mouse cDNA library"
                       /dev_stage="adult"
                       287. .1300
                       /note="unnamed protein product; MO25-LIKE PROTEIN homolog
       CDS
                        [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                       98.2%ID, 100%length, match=1002)
                       putative"
                       /codon_start=1
                        /protein_id="BAC36470.1"
                        /db xref="GI:26345638"
                        /db xref="MGI:1914081"
```

/translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK SLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKK AAP"

BASE COUNT 430 a 294 c 306 g 347 t ORIGIN

Query Ma Best Loc Matches	tch 84.7%; Score 858.8; DB 11; Length 1377; al Similarity 90.4%; Pred. No. 1.3e-175; 917; Conservative 0; Mismatches 97; Indels 0; Gaps	0;
Çy Db	1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC	
Qy Table	61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC	-
	#31 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAACAA 	
Qy	CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG	
	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA	
2y	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	
Ďβ. Ĉè.	361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 	
QΥ	421 CGTTGTGGGAT'IATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	
Db Db	ACT TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	1
Qy Db	541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACATAAAGTGTTGGTAGCAGACATAAAGTGTTGGTAGCAGACATAAAGTATTGGTAGCAGACATAAAGTATTGGTAGCAGACATAAAGTATTGGTAGCAGACATAAAGTATTGGTAGCAGACATAAAGTATTGGTAGCAGACAGA	1
Qy Db	601 TTCTTAGAACAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTGCTTCAGTCTGAGAGACTATGAGAAATTATGACACTATTTTTGAAGACTATGAGAAAATTATGACACTATTTTTGAAGACTATGAGAAAACTGCTGCAATCTGAGAAAATTATGACACTATTTTTGAAGACTATGAGAAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAGAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAGACTATGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAGAACTGCTGCAATCTGAGAGAACTGCTGCAATCTGAGAGAAACTGCTGCAATCTGAGAGAAACTGCTGCAATCTGAGAAACTGCTGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGAAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGCAATCTGAGAAACTGCTGAAAACTGCTGAATCTGAGAAACTGCTGAATCTGAGAAACTGCTGAATCTGAAAACTGCTGAATCTGAAAACTGCTGAAAACTGCTGAAAACTGCTGAAAACTGCTGAAAACTGCTGAAACAAAC	1

```
661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
QУ
            947 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 1006
Db
        721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
ÇУ
            1007 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1066
Db
        781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
QУ
            1067 CTGCTTCGAGACAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1126
Db
        341 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
0.7
            1127 GTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1186
\Gamman
        901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
Qy
            1187 CTCATTGAGTTTTTGAGCAGCTTTCAGAAAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1246
Ūο
        961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Qу
            1247 GAGAAGAACTACCIGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1300
INTSULT S
EK013205
                                       mKNA
                                                        HTC 05-DEC-2002
                               .1449 bp
                                                linear
          AK013205
LOCUS
          Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
NOTTINITEC
          enriched library, clone:2810430N08 product:MO25-LIKE PROTEIN
          homolog [Homo sapiens], full insert sequence.
          AK013205
ACCESSION
          AK013205.1 GI:12850419
VERSION
          HTC; CAP trapper.
KEYWORDS
          Mus musculus (house mouse)
GOURCE
          Mus musculus
  ORGANISM
          Eukaryota; Metazoa; Chordatá; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae: Mus.
REFERENCE
          Carninci, P. and Hayashizaki, Y.
  AUTHORS
          High-efficiency full-length cDNA cloning
  TITLE
          Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
          99279253
  MEDLINE
          10349636
   PUBMED
REFERENCE
          Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  AUTHORS
          Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
          Normalization and subtraction of cap-trapper-selected cDNAs to
  TITLE
          prepare full-length cDNA libraries for rapid discovery of new genes
          Genome Res. 10 (10), 1617-1630 (2000)
  JCURNAL
  MEDLINE
          20499374
  TUPMED
          11042159
REFERENCE
          Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  AUTHORS
          Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
```

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

```
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
            Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
            Ycneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
            Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
 TITLE
            sequencing pipeline with 384 multicapillary sequencer
            Genome Res. 10 (11), 1757-1771 (2000)
  JOURNAL
            20530913
  MEDLINE
            11076861
   PUBMED
REFERENCE
            Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
  AUTHORS
            Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
            Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
            Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
            Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
            Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
            Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
            Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
             Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
             Fletcher, C., Fujita, M., Garibóldi, M., Gustincich, S., Hill, D.,
             Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
             Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
             Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
             Calo, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
             Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
             Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
             and Hayashizaki,Y.
             Junctional annotation of a full-length mouse cDNA collection
  TITLE
             Nature 409 (6821), 685-690 (2001)
  JOURNAL.
             21085660
  MEDLINE
             11217851
   PUBMED
REFERENCE
             The FANTOM Consortium and the RIKEN Genome Exploration Research
   SACHTUA
             Group Phase I & II Team. .
             Analysis of the mouse transcriptome based on functional annotation
   TITLE
             of 60,770 full-length cDNAs
             Nature 420, 563-573 (2002)
   JOURNAL
             6 (bases 1 to 1449)
REFERENCE
             Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
   AUTHORS
             Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
             Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
             Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
              Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
              Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
              Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
              Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
              Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
              Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
              Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
              Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
              Direct Submission
   TITLE
              Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
   JOURNAL
              Physical and Chemical Research (RIKEN), Laboratory for Genome
              Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
```

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

il.

```
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
         URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
         Fax:81-45-503-9216)
         Please visit our web site (http://genome.gsc.riken.go.jp/) for
COMMENT
         further details.
         CDNA library was prepared and sequenced in Mouse Genome
         Encyclopedia Project of Genome Exploration Research Group in Riken
         Genomic Sciences Center and Genome Science Laboratory in RIKEN.
         Division of Experimental Animal Research in Riken contributed to
         prepare mouse tissues. First strand cDNA was primed with a primer
         prepared by using trehalose thermo-activated reverse transcriptase
         and subsequently enriched for full-length by cap-trapper. cDNA went
         through one round of normalization to Rot = 7.5 and subtraction to ...
         Rot = 37.5. Second strand cDNA was prepared with the primer adapter
         of sequence [5'
         with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
         Host: SOLR.
                Location/Qualifiers
PENTURES
                1. .1449
    source
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6J"
                /db xref="FANTOM_DB:2810430N08"
                /db xref="MGI:1893512"
                /dh_xref="taxon:10090"
                /cloue="2310430N08" `
                /tissue_type="whole body"
                /clone_lib="RIKEN full-length enriched mouse cDNA library"
                /dev_stage="10, 11 days embryo"
                281. .1292
       4. ture
                /note="MO25-LIKE PROTEIN homolog [Homo sapiens]
                (SWISSPROT|Q9H9S4, evidence: FASTY, 98.2%ID, 100%length,
                match=1002)
                putative"
                /db xref="MGI:1914081"
                                                    1 1 1 P
                                 367 t
                   304 C
                          325 g
             453 a.
SAST COUNT
OREG.RI
                    83.3%; Score 844.8; DB 11; Length 1449;
  Query Match
  Best Local Similarity 90.2%; Pred. No. 1.4e-172;
                                                     Gaps 1;
                          0; Mismatches
                                       97; Indels
                                                 2;
  Matches 915; Conservative
          1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
OY
           281 ATGAAAAAATGCCCTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 340
Ľ'n
         СĀ
           D_{i}^{t}
        QУ
            Db
        181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
```

Qу

at v

Db		4	161		520
Qy		. 2	241	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA	300
Db1		- 5	521	ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAGAGATGTGACCCAGATA	580
Qγ			301	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	360
Db		į	581	TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT	640
QУ		3	261	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	420
Db	٠.	6	541	TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA	700
Qy		-	21	CGTTGTGGGATTATGCTGAGAAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	480
Db		7	701	CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA	760
ΩУ		Ş	81	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	540
Db		7	76.1	${\tt TTTTCTAATCAGTTCAGAGATTTCTTCAAGTGTTGAGCTGTCCACCTTTGATATCGCT}$	818
Ŏλ		-	eed. T	TCAGATGCCTT'IGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	600+
Db		:	, 4 0 ,	TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC	87:81
Q7				TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	
				TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG	
24 m				ANTTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	
DĐ				AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGG'IGAGCTGATCCTGGACCGCCAC	
Ç7 				AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	
Dĺ.				AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC	
Qу				CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	
Db				CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	
Qy Db				GTGGCCAGCCCCACAAACGCCGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	
				CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC	
Qy Db					
Qγ		5	961	GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014	
Db		1.2	239		

```
AU125107
                                                           EST 01-AUG-2002
                                           mRNA
                                                   linear
                                  822 bp
LOCUS
           AU125107 NT2RM4 Homo sapiens cDNA clone NT2RM4001047 5', mRNA
DEFINITION
           sequence.
           AU125107
ACCESSION
           AU125107.1 GI:10949823
VERSION
           EST.
KEYWORDS
           Homo sapiens (human)
SOURCE
           Homo sapiens
  ORGANISM
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhiņi; Hominidae; Homo.
           1 (bases 1 to 822)
REFERENCE
           Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J.,
  AUTHORS
           Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S. and
           Isoqai, T.
           HRI human cDNA project (Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S.,
  TITLE
           Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki
           Y., Sugano, S., Isogai, T.)
           Unpublished
  JOURNAL
           Contact: Takao Isogai
~COMMENT
           Genomics Laboratory
           Helix Research Institute
           1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
           Tel: 81-438-52-3975
           Fax: 81-438-52-3986
            Mail: genomics@hri.co.jp
           TRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
           Research Institute; cDNA library construction: Department of
           Virology, Institute of Medical Science, University of Tokyo, and
           Belix Research Institute.
                    Location/Qualifiers
AND TURES
                    1. .822
     sobjects:
                    /organism="Homo sapiens"
                    /mol type="mRNA"
                    /db xref="taxon:9606"
                    /clone="NT2RM4001047"
                    /cell type="teratocarcinoma"
                    /cell_line="NT2"
                    /clone lib="NT2RM4"
                    /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
                    precursor cells"
                        164 c 171 g 216 t
                                                   3.others
 PASE COUNT
               268 a
ORIGIN
                         76.0%; Score 770.6; DB 9; Length 822;
  Query Match
                         98.5%; Pred. No. 1.5e-156;
  Best Local Similarity
  Matches 798; Conservative 0; Mismatches
                                              10;
                                                    Indels
           19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
 ÇУ
              1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60
 Db
           79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGCTTCAGAAGAAGTGTCTAAATCACTG 138
 Çу
              .61 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120
 Db
```

ÇУ

Db	121		180
Qу	199	GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG	258
Dp.		GCTCAGCTAGCACAGAACTCTACAGCTGTGACTGTGACACTGACACTGACCTGGCTAGCACAAGAACTCTACAGCAGTGGCCTGGTGACACTGATAGCTGACCTG	
Qy	259	CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA	318
Db		CAGCTGATAGACTTTGAGGGAAAAAAGATGTGTGTGTGTG	
		TO SACRAD CICA CECCA CECCA CECCETA CECTA CECTAGA GENERAL TRATEGE CECCATA CECCATA TECTAGA GENERAL TRATEGE CECCATA CECATA	
ŌΆ		AGACAGATAGGCACTCGGAGTCCTACTGTGGTGTTTTTTTT	
Db		THE TRACETAR AS A CONTATION ACCOCACAGATTGCCTTACGTTGTGGGATTATGCTG	
Qy		TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCTTACGTTGTGGGGATTATGCTG TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGGATTATGCTG	
Db			
CA .		AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA	
ďŒ		AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCMTTTOTAA	
QY	400	GATTTCTTTAAGTACGIGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT	
Dia		JATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT	54.03
Ω γ *	174	TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC	61850
	541		: 600; <u>;</u>
98	ren.	GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA	578.
Dav		GACACTATTTTGAAGACTATGAGAAATTGGTTGAGTGTTTTTGAAGACTATGATGAGAAATTGGTTCAGTCTGAGAAATTATGTTACTAAGAGA 	
		TO GEOGRAPH A A CONTROLLA GEOGRACITE A TECTTOGA CEGTCA CAACTTT GECATCA TOACA	738,
Qy		CAGTCTTTAAAGCTGCTAGGGGAGCTGATCGTGATCATTAAAGCTGTTAAAGCTGCTAGGGGAGCTGATCATGACATCATGACATCTTTGCCATCATGACATCATGACATCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA	
Dρ		PAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT	
QΥ			
Db		1 AAGTATATCAGCAAGCCGGAGAACCTG-AACTCATGATGAACCTNCTTCGGGAT-AAAG	
СY	79	9 CCCAACATCCAGTTTGAAGCCTTTCATGTT 828	
Db	7.7	CCCAACATCCAGTTTGAGCCTTCTGGTTTT 808	•
RESULT 7 BX393735		DX393735 1201 bp mRNA linear EST 13-MA	
LOCUS DEFINITI	ON	DX393735 Homo saviens NEUROBLASTOMA COT 25-NORMALIZED Homo sa	piens
ACCESSIC	N	CDNA clone CS0DC002YI01 5-PRIME, mRNA sequence. BX393735	
VERSION KEYWORDS	;	BX393735.1 GI:30624044 EST.	
SCURCE		Homo sapiens (human)	

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 ORGANISM
         Homo sapiens
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            (bases 1 to 1201)
REFERENCE
         Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
         Full-length cDNA libraries and normalization
 AUTHORS
 TITLE
         Unpublished
 JOURNA'L
          Contact: Genoscope
COMMENT
         Genoscope - Centre National de Sequencage
          BP 191 91006 EVRY cedex - France
          Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
          Library was constructed by Life Technologies, a division of
          Invitrogen. This sequence belongs to sequence cluster 6951.r For
          more information about this cluster, see
          http://www.genoscope.cns.fr/
          cgi-bin/cluster.cgi?seq=CS0DC002AE01QP1&cluster=6951.r. Contact :
          Feng Liang Email : fliang@lifetech.com URL :
          http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
          Faraday Avenue Genoscope sequence ID : CS0DC002AE01QP1.
                 Location/Qualifiers
PEATURES
                 1. .1201
    source
                 organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db xref="taxon:9606"
                  /clone="CS0DC002YI01"
                  /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                  /clone lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                  /note="1st strand cDNA was primed with a NotI-oligo(dT)
                  primer. Five prime end enriched, double-strand cDNA was
                  digested with Not I and cloned into the Not I and EcoR V
                  sites of the pCMVSPORT 6 vector. Library was normalized."
                                            70 others
                            239 g 321 t
                     223 ℃
              348 a
 BASE COUNT
 CRIGIN
                            Score 750.8; DB 13; Length 1201;
                      74.0%;
  Mery Match
                            Pred. No. 3.1e-152;
                      91.0%;
   Post Local Similarity
                                                         Gaps
                                              Indels
                           45; Mismatches
                                          27;
  Matches 766; Conservative
           1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
            ÕΛ
         323 ATGAAAAAATGCCTTTGTTTAGTWAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 382
 Db:
          Qy
         Db
         QУ
          443 GWAGTGTCTWAWTCACTGCTAGCWATGWWAGATATTYTGTGTGGTACAWACGAGWAAGAT 502
 Db
          181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
             CY
          503 CCCCCAACAGAAGCTGTGGCTCAGCTTGCAYAAGWTYTCTTCAGYWGTTGCCTGCTAGTG 562
  Db
          241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
             {{\}}}}
  ÒУ
          563 ACACTGATWGCTGACCTGCAGCTGATAGACTTTKAGGGAAAARDAGATGTGACCCAGATT 622
  aa
```

				260
Qу		TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGT		
Db				
Qу		GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGA		
Db				
Qу		CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAA		
ďQ	743			
Qу	4,81	TTTTCTAATCAAT'ICAGAGATTTCTTTAAGTACGTGGAGTTGTCAACAITTC		
Db				
Oy Db		TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGC		
טט		THE GOLD OF THE PROPERTY AND A CACTATGAGAAATTGCTT	CAGTCTGAG	
Qy DB		TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGACTATGAGACTATGAGACTATTTTTGAAGACTATTTTGAAGACTATTTTGAAGACTATTTTGAAGACTATTTTTTTT	1 1 1 1 1 1 1 1 1	
	- مرجر ر	L AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTA-GGGGAGCTGATCCT	GGACCGTCA	719
. b	·	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTHGGGGGRGCTGATCCT	GGACCGTCA	1042
QУ	. 90	© CAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACT	CATGATGAA	779 ?′
Ola Ola			CMTGATGAA	1102
ूर वेष		0 CCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTT : : : :	• • • • • • • • • • • • • • • • • • • •	I
Ci	. <u>9</u> .4	0 TG 841		3 -
ĞΣ		*[]		· , .
Dh) TL(6	50 TG 1161		
٠		A contract of the contract of		
AF	SULT 8 (005323 CUS SFINITION	Mus musculus adult male cerebellum cDNA, RIKEN full- library, clone:1500031K13 product:MO25-LIKE PROTEIN	HTC 05-DE length en homolog [riched
3.	OCHCCION	sapiens], full insert sequence. AK005323	••	•
	CCESSION ERSION	AK005323.1 GI:12837793		•
	EYWORDS	HTC; CAP trapper. Mus musculus (house mouse)	•	*
	OUTCE ORGANISM EFERENCE	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;	Euteleost Murinae;	omi; Mus.
	AUTHORS	Carninci, P. and Hayashizaki, Y.		

```
High-efficiency full-length cDNA cloning
  TITLE
            Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
            99279253
  MEDLINE
            10349636
   PUBMED
REFERENCE
            Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  AUTHORS
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
  TITLE
            prepare full-length cDNA libraries for rapid discovery of new genes
            Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
             20499374
  MEDLINE
             11042159
   PUBMED
REFERENCE
             Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carnincí, P.,
  AUTHORS
             Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
             Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
             Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
             Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
             Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
             Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
  TITLE
             sequencing pipeline with 384 multicapillary sequencer
             Genome Res. 10 (11), 1757-1771 (2000)
  JOURNAL -
             20530913
  MEDLINE
             11076861
   PUBMED
  THENCE
             Mawai, J. Shinagawa, A., Shibata, K., Yoshino, M., Iton, M., Ashib, Y., Koo
  2200000
             Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Fe
             Alzawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
             Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito R.,
             Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
             Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
             Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
             Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
             Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
             Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
             Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
             Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
             Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
             Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
             Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
             Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
             Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
             Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
             and Hayashizaki, Y.
             Functional annotation of a full-length mouse cDNA collection
   TITLE .
             Mature 409 (6821), 685-690 (2001)
   JOURNA'
             21085660
   MEDLINE
             11217851
    PUBMED
 REFERENCE
             The FANTOM Consortium and the RIKEN Genome Exploration Research
   AUTHORS
             Group Phase I & II Team.
             Analysis of the mouse transcriptome based on functional annotation
   TITLE.
             of 60,770 full-length cDNAs
             Nature 420, 563-573 (2002)
   JOURNAL
                 (bases 1 to 1379)
 REFERENCE
             6.
             Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
```

AUTHORS

```
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
      Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
      Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
      Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
      Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
      Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
      Saitc, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
      Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
      Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
      Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
      Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
      Direct Submission
      Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
      Physical and Chemical Research (RIKEN), Laboratory for Genome
      Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
      RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
      Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
      URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
      Fax:81-45-503-9216)
      Please visit our web site (http://genome.gsc.riken.go.jp/) for
      further details.
      cDNA library was prepared and sequenced in Mouse Genome
      Encyclopedia Project of Genome Exploration Research Group in Riken
      Genomic Sciences Center and Genome Science Laboratory in RIKEN.
      Division of Experimental Animal Research in Riken contributed to
      repare mouse tissues. First strand cDNA was primed with a primer
      greepared by using trehalose thermo-activated reverse transcriptase \mathbf{p}_{AB}
      and subsequently enriched for full-length by cap-trapper. Second
      Strand cDNA was prepared with the primer adapter of sequence[5'
      was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3'
       end: SstI. Host: SOLR.
               Location/Qualifiers
               1. .1379
Dource
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6J"
                /db_xref="FANTOM_DB:1500031K13"
                /db xref="MGI:1901050"
                /db xref="taxon:10090"
                /clone="1500031K13"
                /sex="male"
                /tissue_type="cerebellum"
                /clone_lib="RIKEN full-length enriched mouse cDNA library"
                /dev stage="adult"
                285. .1175
                /note="unnamed protein product; MO25-LIKE PROTEIN homolog
                [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                98.2%ID, 100%length, match=1002)
                putative"
                /codon_start=1
                /protein_id="BAB23953.2"
                /db_xref="GI:26342524"
                /db xref="MGI:1916258"
                /translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
```

PLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN

TITLE

COMMENT

M. URES

JCURNAL

NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDT1FEDYEKLLQ SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV

FKNSVFITNRIHGLKRWLSS" 1361. .1366 polyA_signal /note="putative"

/note="putative" 284 C 291 g 452 a

1379

polyA_site

352 t BASE COUNT ORIGIN

Query Match 70.0%; Score 709.4; DB 11; Length 1379; Best Local Similarity 88.9%; Pred. No. 3e-143; Matches 767; Conservative 0; Mismatches 96; Indels 0; Gaps 0;	
Qy 1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60	
Cy 51 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC	
QY GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	
CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTAGTG 240:	
241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300	
Db 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360	
27 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420	
QY 421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480	
Qy 481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540	Į.
QY 561 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600	1
Qy 60. TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660)

```
885 TTCTTAGAACAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 944
Db
          661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
              QУ
          945 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 1004
Db
          721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
              41 | 16 | 414 | 414 | 444 | 444 | 444 | 444 | 444 | 444 | 444 | 444 | 444 | 444 | 444 | 444 | 444 | 444 | 444 |
QУ
         1005 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1064
Db
          781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
            Qу
         1065 CTGCTTCGAGACAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGAATTCT 1124
Db
          841 GTGGCCAGTCCTCACAAAACACA 863
QУ
                 1125 GTCTTTATTACAAATAGAATACA 1147
aC.
RESULT 9
DG218735
                                                               EST 21-APR-2001
                                                       linear
                                    784 bp mRNA
            BG218735
            RST38476 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
LOCUS
 DEFINITION
            3G218735
 ACCESSION
            BG218735.1 GI:13744756
 VERSION
            EST.
 KUKWORDS
            Homo sapiens (human)
 CURCE
                                                                    1 /3-1-
            Homo sapiens
   MISM
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
             : (bases 1 to 784)
           Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
 CRENCE
            Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.
   AUTHORS
             , Lerner, L., Costanzo, D., McElligott, K., Boozer, S., Mays, R., Smith
             ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
             ,J., Danzig,J. and Ducar,M.
            Creation of genome-wide protein expression libraries using random
             activation of gene expression
             Nat. Biotechnol. 19 (5), 440-445 (2001)
   JUURNAL
             21227151
   MEDLINE
             11329013
    PURMED
                                                                      4 · \( \xi_1 \)
             Contact: Scott J. Cain
 COMMINT
             Athersys, Inc.
             3201 Carnegie Ave, Cleveland, OH 44115, USA
             Tel: 216 431 9900
             Fax: 216 361 9596
             Email: scain@athersys.com
             High quality sequence stop: 515.
                      Location/Qualifiers
  FEATURES.
                      1. .784
      source
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /cell_line="HT1080"
                      /clone_lib="Athersys RAGE Library"
                      /note="See 'Creation of Genome-wide Protein Expression
                      Libraries using Random Activation of Gene Expression',
```

Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

BASE COUNT

265 a 151 c 158 g 209 t 1 others

Query Matc	h Similarity	66.3%;	Score	671.8;	DB 10;	Length	784;		
Best Local Matches: 7	18; Conservat	cive	0; Mis	matches	23;	Indels	3;	Gaps	. 3;
ОУ	1 ATGAAAAAAAT	GCTTTGT	TTAGTA	AAŤCACACA !	AAAATCO	CAGCAGAAA	TTGTG	AAAATC	60
Db 3	2. ATGAAAAGAAT	GCCTTTGT	TTAGTA	AATCACACA	AAAATC	CAGCAGAAA	ATTGTG'	TAAATC	91
	3. CTGAAAGACAA'	111111		:	.				
Db 9	2 CTGAAAGACAA	TTTGGCCA	TTTTGG	AAAGCAAC	SACAAAA	AGACAGACA	\AGGCT	TCAGAA	151
	1 GAAGTGTCTAA		11111			111111		11111	
	2 GAAGTGTCTAA	ATCACTGC	'AAGCAA'	TGAAAGAA!	ATTCTGT(GTGGTACA/	AACGAG	AAAGAA	
	CCCCAACAGA	1111111	1 1111.	1	!	111111	11111		
	2 CCCCCAACAGA	AGCAGTGG	CCCAGC	TAGCACAA(JAACTCT.	ACAGCAGT'	3GCCT6	CTAGTG	
****	ACACTGATAGC	11111111		1111111		11111		11111	
0.5°	•			•					
⁷ Ωγ	TTTAACAACAT	11111111111	111111	1111111		111111	11111		*
	T GCTCATCCTCA							•	
4.24	GT GCTCATCCTCA 	1111111			111111	11111	11111	.	
	21 CGTTCTGGGAT								
~ 2	21 CGTTCTGGGA 52 CGTTGTGGGAT		11111	1111111	1111111				
C γ 4	81 TTTTCTAATC	ATTCAGA(GATTTCT	TTTAAGTAC	GTGGAGT	TGTCAACA	TTTGAT	TATTGCT	54·0°
	1.2 TTTTCTAATCA	AATTCAGA	 GATTTCT	 TTAAGTAC	GTGGAGT	TGTCAACA	 TTTGAT		571.
Çy 5	4.1 TCAGATGCCT	TTGCTACT	TTCAAGO	SATTTACTA	ACCAGA	CATAAAGTG	TTGGT	AGCAGAC	600%
pb 5	 72 TCAGATGCCT	 TGCTACT	 TTCAAGO	ATTTACTA	ACCAGAC	CATAAAGTO	TTGGT	AGCAGAC	631
Qγ 6	01 TTCTTAGAACA	AAAATTAC	GACACTA	4. 	GACTATO	G-AGAAATT	GCTTC	AGTCTGA	659
Db 6	 32 TTCTTAAACA	IIII AAATTACG	ACACTA	TTTTTGA <i>F</i>	GACTAT	GAAGAAATT	GCTTC	AGTCTGA	691 [.]
Qγ €	60 GAATTATGTT	AC-TAAGA 	GACAGT(CTTTAAAGO	CTGCTAGG	EGGAGCTGA	ATCCTG	GACCGTC	718

```
692 GAATTATGTTACTTAAGAGACAGTCTTTAGAGCTGCTAGGGGAGCTGATCCTGAAANGTT 751
Db
          719 ACAACTTTGCCATCATGACAAAGT 742
Q۷
               752 ACAACTTTGCC-TCATGACAAAGT 774
Db
RESULT 1.0
AK013161
                                                                   HTC 05-DEC-2002
                                                          linear
                                      1281 bp
                                                 mRNA
            Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
LOCUS
            enriched library, clone:2810425013 product:MO25-LIKE PROTEIN
DEFINITION
            homolog [Homo sapiens], full insert sequence.
             AK013161
ACCESSION
             AK013161.1 GI:12850350
VERSION
             HTC; CAP trapper.
KEYWORDS
             Mus musculus (house mouse)
SOURCE
             Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  ORGANISM
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
             Carninci, P. and Hayashizaki, Y.
   AUTHORS
             High-efficiency full-length cDNA cloning
   TITLE
             Meth. Enzymol. 303, 19-44 (1999)
   JOURNAL
             99279253
   MEDLINE
             10349636
    PUBMED
             Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., CR
 REFERENCE
              Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
   RUTHORS
             Normalization and subtraction of cap-trapper-selected cDNAssito
              prepare full-length cDNA libraries for rapid discovery of new genes.
   TITLE S
              Genome Res. 10 (10), 1617-1630 (2000)
   COURNAL
              20499374
   MEDLINE
              11042159
   UBMED
              Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 DEPERENCE
              Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
   JUTHORS
              Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
              Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
              Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
              Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
              Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
              RIKEN integrated sequence analysis (RISA) system--384-format
              sequencing pipeline with 384 multicapillary sequencer
    TITLE
              Genome Res. 10 (11), 1757-1771 (2000)
    JOURNAL
              20530913
    MEDLINE
              11076861
     PUBMED
              Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
  REFERENCE
              Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
    AUTHORS
              Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
               Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
               Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
               Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
               Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
               Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
               Wagner, L., Wasnio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
```

Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki,Y. Functional annotation of a full-length mouse cDNA collection TITLE Nature 409 (6821), 685-690 (2001) JOURNAL 21085660 MEDLINE 11217851 PUBMED REFERENCE The FANTOM Consortium and the RIKEN Genome Exploration Research **AUTHORS** Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation TITLE of 60,770 full-length cDNAs Nature 420, 563-573 (2002) JOURNAL (bases 1 to 1281) REFERENCE Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., AUTHORS Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Giraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Lasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda. M., Hoya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Momura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., ⊱ızuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y. Direct Submission TITLE Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of JUSPNAL Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216) Please visit our web site (http://genome.gsc.riken.go.jp/) for COMMENT further details. CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA wentthrough one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR.

```
Location/Qualifiers
FEATURES
               1. .1281
   source
               /organism="Mus musculus"
               /mol type="mRNA"
               /strain="C57BL/6J"
               /db_xref="FANTOM_DB:2810425013"
               /db_xref="MGI:1908997"
               /db xref="taxon:10090"
               /clone="2810425013"
                /tissue_type="whole body"
               /clone_lib="RIKEN full-length enriched mouse cDNA library"
                /dev stage="10, 11 days embryo"
                289. .1096
   misc feature
                /note="MO25-LIKE PROTEIN homolog [Homo sapiens]
                (SWISSPROT | Q9H9S4, evidence: FASTY, 98.2%ID, 100%length,
                match=1002)
                putative"
                /db xref="MGI:1922871"
                                332 t
                   275 C
                          291 g
            333 a
BASE COUNT
ORIGIN
                                          Length 1281;
                          Score 622.8; DB 11;
                    61.4%;
  Query Match
                          Pred. No. 1.7e-124;
                    82.8%;
 Best Local Similarity
                                                          2:
                                      97; Indels
                                                62;
                                                    Gaps
                            Mismatches
                         0:
 Matches 763; Conservative
        Qy
           1)57
        153 AATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTGGCTCAGCTAGCACA 212
           . . . . . .
        AATTCTGTGTGGAACGAACGACAAGGAGCCCCCTACAGAAGCAGTGGCTCAGCTGGCGCA 35690
100
        3 AGAACTCTACAGCAGTGGCCTAGTGACACTGATAGCTGACCTGCAGCTGATAGACTT 272
                         357 GGAGCTCTACAGCAGCGGGTTGCTGGTGACACTCATAGCTGACCTGCAGCTCATAGACTT 416
IJ.)
        273 TGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGACAGATAGGCAC 332
           Qŷ
        417 TGAGGCAAAAAAAGATGTGACCCAGATATTCAACAACATCCTGAGAAGACAGATTGGTAC 476
Db
        233 TCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAA 392
            Qy.
        477 ACGGTGTCCTACTGTCGAGTACATCAGTTCTCATCCTCACATCCTGTTTATGCTTCTCAA 536
 Db
        323 AGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCG 452
           Çγ
        537 AGGCTATGAAGCCCCACAGATTGCCTTACGCTGTGGGATTATGCTAAGAGAGTGTATTCG 596
 Db.
        453 ACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTA 512
           97
        597 ACATGAGCCACTTGCCAAAATCATCCTATTTTCTAATCAGTTCAGAGATTTCTTCAAGTA 656
 Db
        513 CGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACT 572
 Qу
            657 TGTTGAGCTGTCCACCTTTGATATCGCTTCAGATGCCTTCGCTACTTTTAAG----- 708
 Db
```

```
573 AACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGA 632
Qу
                                                        -ATTTTTGA 716
Db
        63.3 AGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCT 692
Qу
            717 AGACTATGAGAAACTGCTGCAATCTGAGAACTATGTGACAAAGAGACAATCTTTAAAGTT 776
Db
        693 GCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAA 752
QУ
            777 GCTAGGTGAGCTGATCCTGGACCGCCACAATTTCACCATTATGACCAAG--TATCAGCAA 834
Db
        753 GCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTT 812-
Qy
            335 GCCAGAGAACCTGAAACTGATGATGAACCTGCTTCGAGACAAAAGTCCCAACATCCAATT 894
Db
        813 TGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGT 872
Qy
             $95 CGAAGCCTTCCATGTCTTTAAGGTGTTTGTGGCCAGCCCCCACAAAACGCAGCCTATCGT 954
Db
        873 GGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGA 932
QУ
            955 GGAGATTCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTTCAGAAAGA 1014
Dio
        933 AAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGA 992-
ΩУ
           _015 AAGGACAGACGACGAGCAGTTTGCTGACGAGAAGAACTACCTGATTAAACAGATTCGAGA 1074
Dio
        993 CTTGAAGAAAACGGCCCCTTGA 1014
ÇV
            1075 CTTGAAGAAAGCAGCCCCGTGA 1096
RESULT 11
BU116523
                                               linear
                                                      EST 25-NOV-2002
                                       mRNA
                               951 bp
LOCUS
          BU116522
          603139786F1 CSEQCHL15 Gallus gallus cDNA clone ChEST129122 5', mRNA
DEFINITION
          sequence.
          £U116522
ACCESSION
          BU116522.1 GI:25323402
VERSION
KEYWORDS
          EST.
          Gallus gallus (chicken)
SOURCE
  ORGANISM
          Gallus gallus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
          Phasianinae; Gallus.
             (bases 1 to 951)
REFERENCE
          Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
  AUTHORS
          Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
          A Comprehensive Collection of Chicken cDNAs
  TITLE
          Curr. Biol. 12 (22), 1965-1969 (2002)
  JOURNAL
  MEDLINE
           22335534
           12445392
   PUBMED
           Contact: Simon Hubbard
COMMENT
           Department of Biomolecular Sciences
          University of Manchester Institute of Science and Technology (UMIST
```

```
PO Box 88, Manchester, M60 1QD, UK
         Tel: 01612008930
         Fax: 01612360409
         Email: Simon. Hubbard@umist.ac.uk.
                 Location/Qualifiers
FEATURES
                 1. .951
   source
                 /organism="Gallus gallus"
                 /mol type="mRNA"
                 /strain="Compton Line 15I"
                 /db xref="taxon:9031"
                 /clone="ChEST129122"
                 /sex="Female"
                 /tissue type="cerebrum"
                 /dev stage="adult"
                 /lab host="DH10B"
                 /clone_lib="CSEQCHL15"
                 /note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
                 EcoRI; Site_2: Not1; Modification of pBluescript II KS(+)
                 [Stratagene] vector to accommodate cDNA produced with the
                 T-trimmed protocol (Construction of uni-directionally
                 cloned cDNA libraries from messenger RNA for improved 3'
                 end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387
                 ,624). Cut pBluescript II KS(+) with NotI and EccRI.
                 Ligate in double stranded adaptor containing BsgI and
                 BamHI sites [5'ggccgcgtgcagccccggatccgaaaaaaag]
                 [5'eattotttttttcggatccggggctgcacgc]"
                                                         19901 183
                            202 q 240 t
                    305 c
             303 a
BASE COUNC
ORTGIN
                            Score 614; DB 13; Length 951;
                     60.6%;
 Query Match
                                                         e mer.
                     32.8%; Pred. No. 1.3e-122;
 Best Local Similarity
                                                         Gaps
                           0: Mismatches 150; Indels
 Matches 737; Conservative
        \Omega V
           6 GAAGTGTCAAAATCTCTGCAAGCAATGAAGTAAATTCTGTGTGGGACCACAGACAAGGAG 65
181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
QУ
            66 CCACCGACAGAAGTAGTGGCTCAGCTGGCACAAGAATTGTACAACAGTGGCCTTCTAGTG 125
ďC
        241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
ŲУ
            126 ACACTTATTGCCAACCTGCAGCTCATAGATTTTGAGGGTAAAAAGGATGTTTCCCAGATA 185
ĎЬ
        301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Ωy
            186 TTTAACAACATCCTGAGAAGACAAATTGGCACACGAAGCCCTACTGTGGAATACATTAGT 245
Dρ,
        351 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
QУ.
            346 GCCCATCCACATATCCTGTTCATGCTTCTGAAAGGCTATGAATCCCCAAATATTGCCTTA 305
Db
        421 CGTTGTGGGATTATGC'ÍGAGAGAATGTATTCGACATGAACCACTTGCCAA-AATCATCCT 479
QУ
            306 CGCTGTGGAATTATGCTGAGGGAGTGCATCCGACATGAACCATTGGCCAACAATCATACT 365
```

nb

Qy	480 CTTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGC 539
Db	366 TTTTTCAGAACAGTTCAGAGACTTCTTCAAGTATGTGGAAATGTCAACATTTGATATAGC 425
QY	540 TTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGA 599
Db	426 ATCTGATGCCTTTGCTACATTCAAGGACTTGTTAACAAGGCACAAG'TTGTTGGTAGCAGA 485
ОХ	600 CTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGA 659
Db _.	486_TTTTATGGAACAAAATTATGATACGATCTTTGAGGATTATGAAAAACTCCTTCATTCTGA 545
Qу	660 GAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCA 719
Db	
Qy	720 CAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAA 779
Db	
Qу	780 CCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTT 839
Db ·	
Çy -	©40 BUTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA 899
7.5	
17 1 - 12 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	### ##################################
D\$:	
ОУ	959 ACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCC 1008
Db	
RESULT 12	
LOCUS DEFINITIO	BQ669953 982 bp mRNA linear EST 15-JUL-2002 NAME AGENCOURT_8203755 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6255924
	5', mRNA sequence.
ACCESSION	I BQ669953 3Q669953.1 GI:21780787
VERSION KEYWORDS	30009955 till 31.21700707
SOURCE	Homo sapiens (human)
ORGANIS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE AUTHORS	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL COMMENT	. Unpublished Contact: Robert Strausberg, Ph.D.
COMMENT	Email: cgapbs-r@mail.nih.gov
	Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory
	CDMA DIDEATY FICPARACION. RUDIN HABOLACOLY

```
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLCM2407 row: m column: 13
          High quality sequence stop: 508.
                 Location/Qualifiers
FEATURES
                  1. .982
    source
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6255924"
                  /tissue_type="epidermoid carcinoma, cell line"
                  /lab host="DH10B (phage-resistant)"
                  /clone_lib="NIH_MGC_102"
                  /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
                  Site_2: EcoRI; cDNA made by oligo-dT priming.
                  Directionally cloned into EcoRI/XhoI sites using the
                  following 5' adaptor: GGCACGAG(G). Library constructed
                  by Ling Hong in the laboratory of Gerald M. Rubin
                  (University of California, Berkeley) using ZAP-cDNA
                  synthesis kit (Stratagene) and Superscript II RT (Life
                  Technologies). Note: this is a NIH MGC Library."
                                             11 others
                                     197 t
                             357 q
                      200 C
              217 a
BASE COUNT
ORIGIN
                             Score 594.2; DB 13; Length 982;
                      58.68;
  Guery Match
                             Pred. No. 2.6e-118;
                      96.5%;
  Dest Local Similarity
                             0; Mismatches
                                           20;
                                               Indels
  Maydhes 623; Conservative
         334 CGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAA 393
1 CGGAGTCCTACTGTGGAG-ATATTAGTGCTCATCCTCATATCCTGGTTATGCTCCTCAAA 59
         394 GGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGA 453
            50 GGATATGAAGCCCCACAGATTGCCTTACATTGGGGGATTATGCTGAGAGAATGGATTCGA 119
מני
         454 CATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTAC 513
             QY
         220 CATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTAC 179
\mathcal{D}_{i}
         514 GTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTA 573
 Ç
             180 GTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTA 239
 Dis
         574 ACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAA 633
             Qу
         240 ACCAGACATAAAGTGGTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAA 299
 Db
         534 GACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTG 693
             ************************************
 Qy
         300 GACTATGAGAAATTGCTTCAGTCTGAGAATTATGGTACTAAGAGACAGTCTTTAAAGCTG 359
 Db
         694 CTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAG 753
             QY
```

```
360 CTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAG 419
Db
         754 CCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTT 813
             QУ
         420 CCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTT 479
Db
         814 GAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTG 873
             QУ
         480 GAAGCCTTTCATGGTTTTAAGGGGGTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTG 539
Db
         874 GAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAA 933
             QУ
         540 GAGATCCTGGTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAAGAA 599
Db
         934 AGG--ACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAAC 982
                    BRITANIA II HIIMBEL TITTI I III
QY
         600 AGGGACGGGATGATGAGCANNTCCCTGACGAGAAAGACTACTTGGGTTAAC 650
Db
RESULT 13
3U518807
                                                            EST 12-SEP-2002
                                                   linear
                                            mRNA
                                   934 bp
           BU518807
LOCUS
           AGENCOURT_10171930 NIH_MGC_134 Mus musculus cDNA clone
DEFINITION
           IMAGE: 6516567 5', mRNA sequence.
           BU518807
ACCESSION
           BU518807.1 GI:22826333
752SION
                                                               KETTIE
           ST.
  ETWORDS
                                                              - BOURTS
           Fus musculus (house mouse)
 BUTT OF
                                                                09655
  CARANISM Was musculus
           mukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            (bases 1 to 934)
 もこでは変数NCE
            WIH-MGC http://mgc.nci.nih.gov/.
  AUTHORS
            Mational Institutes of Health, Mammalian Gene Collection (MGC)
  TITLE
            Unpublished
   COURNAL
                                                                17 MM 12 1
            Contact: Robert Strausberg, Ph.D.
2000ENT
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. David Rowe
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14095 row: e column: 16
            High quality sequence stop: 656.
                    Location/Qualifiers
 FEATURES
                     1. .934
      source
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6516567"
                     /tissue_type="undifferentiated limb"
                     /lab host="DH10B (phage-resistant)"
                     /clone lib="NIH_MGC 134"
                     /note="Vector: pCMV-SPORT6.1.ccdb; Site_1: EcoRV; Site_2:
                     NotI; Cloned unidirectionally. Primer: Oligo dT. Average
```

insert size 1.7 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library." 301 a 198 c 200 g 234 t 1 others

BASE COUNT

Бe	ery Match st Local S tches 654	imilarity	57.7%; 89.0%; tive	Pred.	585.2; No. 2.3e	-116;	Length Indels	934; 2;	Gaps	2;
ОУ	`	TGAAAAAAATG TGAAAAAAATG	1111111	411111	111111	11111	1	1111	!	
QУ	62	TGAAAGACAAT TGAAAGACAAC	TTGGCCAT	TTTTGGAP	AAGCAAG <i>I</i>	CAAAAA 	GACAGACAA	GGCTT	CAGAAG	121
ΩУ	122	AAGTGTCTAAA AGGTGTCAAAA	TCACTGC	AAGCAATO	GAAAGAAAT	TCTGTG	TGGTACAAA	CGAGA	AAGAAC	181
Db Qу	182	CCCCAACAGAA	AGCAGTGG(CTCAGCT?	AGCACAAGA	AACTCTA	ÇAGCAGTG(GCTGC	TAGTGA	241
(3) (3)	242	CCCCTACAGAA CACTGATAGCT	rgacctgc	AGCTGAT	AGACTTTGA	AGGGAAA	AAAAGATG	rgacco	AGATAT	301
DZ. Qy⊸	2.12	CACTCATAGCT	CTTGAGAA	GACAGATA	AGGCACTC	GGAGTCC	TACTGTGG	AGTATA	TTAGTG	36403
	等原理 类	TCAACAACATO	PATCCTGT	TTATGCT	CCTCAAAG	GATATGA	AGCCCCAC	AGATTG	CCTTAC	121 2
Óλ	22	CTCATCCTCA(GTTGTGGGAT	TATGCTGA	GAGAATG	TATTCGAC	ATGAACC	CACTTGCCA	AĄATCA 	TCCTCT	481
Db Ç∵	482	GCTGTGGGAT' TTTCTAATCA	ATTCAGAG	ATTTCTT	TAAGTACG	TGGAGTI	GTCAACAT	TTGATA	ATTGCTT	541
Db Qy		TTTCTAATCA								
Ωb		CAGATGCCTT								
Db	713	TCTTAGAACA ATTATGTTAC		 ACACTAT	 TTTTGAAC	 GACTATGA	 AGAAACTGC	 TGNCAT		772
Qу Db		ATTATGTTAC		$ \cdot \cdot \cdot $	4		-GAGCTGAT IGAGCTGAT	\Box		

```
720 CAACTTTGCCATCAT 734
Qу
                   833 CACAATTTTCACCAT 847
Db
```

RESULT 14 CD354831

linear EST 29-MAY-2003 713 bp mRNA CD354831 LOCUS.

UI-M-GMO-cge-i-10-0-UI.rl NIH_BMAP GMO Mus musculus cDNA clone DEFINITION

IMAGE: 30361641 5', mRNA sequence.

CD354831 ACCESSION

CD354831.1 GI:31147332 VERSION

KEYWORDS EST.

TES-

MOUNTAIN

Mus musculus (house mouse) SOURCE

Mus musculus ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 713) REFERENCE

NIH-MGC http://mgc.nci.nih.gov/. AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC) TITLE

Unpublished JOURNAL

Contact: Robert Strausberg, Ph.D. COMMENT

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Towa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Distribution information can be found at

http://genome.uiowa.edu/distribution/mousefl.html

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seg primer: pYX-5.

Location/Qualifiers

1. .713

/organism="Mus musculus"

/mol type="mRNA" /strain="C57BL/6"

/db xref="taxon:10090"

/clone="IMAGE: 30361641"

/tissue type="whole brain"

/dev_stage="1, 5 and 15 days newborn"

/lab_host="DH10B (T1 phage resistant)"

/clone lib="NIH BMA? GMO"

/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site.Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAACTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator."

BASE COUNT 220 a 159 c 140 g 192 t 2 others
ORIGIN

Ouery Match 57.1%; Score 579.2; DB 14; Length 713;

Query Match Pred. No. 4.4e-115; 89.9%; Best Local Similarity Mismatches Indels Gaps 2: Matches 642; Conservative 0; 282 AAAAGATGTGACCCAGATATTTAACAACAT-CTTGAGAAGACAGATAGGCACTCGGAGTC 340 Qy 1 AAAAGATGTGACCCAGATATTCAACAACATCCNTGAGAAGACAGATTGGTACACGGTGTC 60 Db' 341 CTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATG 400 QYCTACTGTCGAGTACATCAGTTCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATG 120 Db 401 AAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAAC 460 ÇУ 121 AAGCCCCACAGATTGCCTTACGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGC 180 Db 461 CACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGT 520 Őλ 131 CACTTGCCAAAATCATCCTATTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGC 240 קנו 501. TGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTÁCTAACCAGAC 580. 1 2 1 TGTCCACCTTGATATCGCTTCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGAC 300, E31 ATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATG 640 ATAAAGTATTGGTAGCAGACTTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATG 360 70% AGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGG 700) \bigcirc AGAAACTGCTGCAATCTGAGAACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTG 420 Db 701 AGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGA 760 Qу 421 AGCTGATCCTGGACCGCCACAATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGA 480 Db 761 ACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCT 820 Qу 481 ACCTGAAACTGATGATGAACCTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCT 540 Db821 TTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCC 880 Qу 541 TCCATGTCTTTAAGGTGTTTGTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTC 500 Db. Qу Db 941 ATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACT 994 ŌХ 561 ACGACGAGCAGTTTGCTGACGAGAAG-ACTACCTGATTANACAGATTCGAGACT 713 Db

```
RESULT 15
HSM073180
     HSM073180 standard; RNA; EST; 742 BP.
XX
AC
     BX483012;
XX
     BX483012.1
SV
ХX
     09-MAY-2003 (Rel. 75, Created)
DT
     09-MAY-2003 (Rel. 75, Last updated, Version 1)
DT
XX
     Homo sapiens mRNA; EST DKFZp686C08234_r1 (from clone DKFZp686C08234)
DΕ
XX
     EST; expressed sequence tag.
KW
XX
     Homo sapiens (human)
OS
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC
     Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC
XX
KN
RP
     1..742
     Ottenwaelder B., Obermaier B., Deutschenbaur S., Mewes H.W., Weil B.,
RA
     Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RA
RT
     Submitted (07-MAY-2003) to the EMBL/GenBank/DDBJ databases.
^{\mathrm{RL}}
     MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
ŔЬ
Ċ!
     This is the 5' sequence of the clone insert
     Clone from S. Wiemann, Mclecular Genome Analysis, German Cancer
CC
    Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC -
     sequenced by MediGenomix (Martinsried/Germany) within the cDNA
CC
     sequencing consortium of the German Genome Project.
CC.
1777
    Mo sl sequence available.
     This clone (DKFZp686C08234) is available at the RZPD in Berlin:
CC
CC
     Please contact the RZPD: Ressourcenzentrum, Heubnerweg 5,
     14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
CC
XX
FH ·
     Кеу
                     Location/Qualifiers
ΞH
E_{\rm eff}
     source
                     1. .742
                     /db xref="taxon:9606"
FT
                     /mol type="mRNA"
FT
                     /organism="Homo sapiens"
                     /clone="DKFZp686C08234"
FT
                     /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
Sir.
                     DH10B; sites SfiIA + SfiIB"
1715
                     /dev stage="adult"
FT
                     /tissue_type="cDNA-collection"
FT
ΧX
     Sequence 742 BP; 256 A; 143 C; 162 G; 179 T; 2 other;
SQ
                          57.0%; Score 578; DB 2; Length 742;
  Query Match
                          99.8%; Pred. No. 8e-115;
  Best Local Similarity
  Matches 578; Conservative.
                                 0; Mismatches
                                                   1;
                                                        Indels
            1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Qу
```

Db	164 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 223
Qy	61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAGACAGAC
Db	
QУ	121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA
Db	
Öλ	181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Db	
Qy Db	2#1 ACACIGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
QΥ	301. TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Dр	
ΟУ	361 SCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
100	
f e e e-si	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480222222222222222222222222222222222222
10.11	
, Q;;:	681 PTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
350	
(// 	541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGA 579
ונע	•••

Search completed: January 5, 2004, 03:18:09 Cob time : 2589 secs

GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model January 6, 2004, 00:37:47; Search time 3965 Seconds Run on: (without alignments) 10462.134 Million cell updates/sec US-10-088-872-1 Title: Perfect score: 1014 l atgaaaaaaatgcctttgtt.....tgaagaaaacggccccttga 1014 Sequence: Scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1.0 2888711 seqs, 20454813386 residues Searched: Total number of hits satisfying chosen parameters: 5777422 Minimum DB seq length: 0 Maximum DB seq length: 2000000000 Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries teat at s GenEmbl:* lydeabase : 1: gb ba:* 2: gb_htg:* 3: gb_in:* 4: gb_om:* 5: gb_ov:* 6: gb_pat:* 7: gb ph:* 8: gb_pl:* 9: gb_pr:* 10: gb_ro:* 11: gb_sts:* 12: gb_sy:* 13: gb_un:*

14: gb_vi:* 15: em_ba:* 16: em_fun:* 17: em_hum:* 18: em_in:* em_mu:* 19: em om:* 20: 21: em_or:* 22: em_ov:* 23: em_pat:* 24: em_ph:* 25: em_pl:* 26: em ro:* 27: em_sts:*

```
28:
    em_un:*
29: em_vi:*
30: em_htg_hum:*
    em_htg_inv:*
31:
    em htg_other:*
32:
    em htg_mus:*
33:
     em_htg_pln:*
34:
     em_htg_rod:*
35:
     em_htg_mam:*
36:
     em_htg_vrt:*
37:
38:
     em_sy:*
     em_htgo_hum:*
39:
     em_htgo_mus:*
40:
     em htgo_other: *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

			•	б					_
Re	esu.	lt,		Query					Description
•		o .	Score	Match	Length	ĎΒ	ID	•	Description
		., -			_ 				AX105381 Sequence
		1	1014	100.0	1014	6	AX105381		AR097361 Sequence
		2	1010.8	99.7	1344	6			AR203365 Sequence
		3	1010.8	99.7	1344	.6	AR203365		BC010993 Homo sapi.
		4	1009.2	99.5	1491	9.	BC010993		BD157871 Primer fo
		Š	992.8	97.9	2002	ti.	BD157871	•	AK022639 Homo sapi
		6	992.8	97.9	2002	9	AK022639		BC016128 Mus muscu
		7	360.4	84.9	1359	10	BC016128		BC016546 Mus muscu
		. દુ	858.8	84.7	1530	10	BC016546		BD147463 Primer to
		3	770.6	76.0	822	. 6	BD147463		BD079551 Cancer-as
		10	Q04.5	67.5	831	6	BD079551		
		11	32.6	57.5	1026	6	AX061831		AX051831 Sequence
		12	582.6	57.5	3281	6	AX082322		AX082322 Sequence
		13	582.6	57.5	3761	9	BC020570		BC020570 Homo sapi
		14	581.6	57.4	1680	9	AF151824		AF151824 Homo sapi
		15	5816	57.4	3466	9	AF113536		AF113536 Homo sapi
		16	568.8	56.1	1947	10	BC020041		BC020041 Mus muscu
		17	567.2	55.9	2322	10			S51858 MO25 gene [
		18	552.8	54.5	3346	9	AK000804		AK000804 Homo sapi
		19	541.6	53.4	1053	б	AX105727		AX105727 Sequence
		20	533.6	52.6	851	9	AY211923		AY211923 Homo sapi
	C	21		51.3	833	6	BD079552		BD079552 Cancer-as
	C	.2.2	478.8	47.2	2366	5	BC044172		BC044172 Danio rer
		23	3,68.8	36.4	2991	10	BC029053		BC029053 Mus muscu
		24	365	36.0	158599	2	AC101941		AC101941 Mus muscu
		25	362.8	35.8	1894	3	AY128438		AY128438 Drosophil
		26	362.8	35.8	8324	2	AC019926		AC019926 Drosophil
	_	27	362.8	35.8	201313	3	AC010688		AC010688 Drosophil
	c c	28	362.8	35.8			AE003526		AE003526 Drosophil
	C	29	361.2	35.6			AB000402		AB000402 Drosophil
		30	270	26.6			AK026335		AK026335 Homo sapi
		31	255	25.1			CPR238632		AJ238632 Chlorella
		32	240.8	23.7			BC034159		BC034159 Mus muscu
		33	210.4	20.7			AY088359		AY088359 Arabidops
		23	210.4	20.7					•

0 0 0 0 0 0 0	34 35 36 37 38 39 40 41 42 43 44 45	209.6 200.2 200.2 195 195 186.4 185.4 185.4 172.8 169.8 166.6		1570 1032 1519 387 387 89137 89807 349980 349980 159863 722 700	8 8 8 6 6 8 9 6 6 9 6 6	AF428296 AY055792 AF380659 AX333971 AX407834 AC006284 AL138875 AX711879 AX739961 AL136218 AX283846 AX283847	AF428296 Arabidops AY055792 Arabidops AF380659 Arabidops AX333971 Sequence AX407834 Sequence AC006284 Arabidops AL138875 Human DNA AX711879 Sequence AX739961 Sequence AL136218 Human DNA AX283846 Sequence AX283847 Sequence
---------------	--	---	--	--	-------------------------	--	---

ALIGNMENTS

```
RESULT 1
                                                                   PAT 30-APR-2001
AX105381
                                                          linear
                                                 DNA
                                      1014 bp
            AX105381
LCCUS
            Sequence 1 from Patent W00123552.
DEFINITION
            AX105381
ACCESSION
            AX105381.1 GI:13921508
VERSION
KEYWORDS
             Homo sapiens (human)
SOURCE
                                                                        3 - 11 1 mo 14:
  ORCANISM . Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
             den Daas, I. and Duecker, K.
             Human paralogue of a head trauma induced cytoplasmatic dalcium
  AUTHORS
  TILE
             binding protein
             Patent: WO 0123552-A 1 05-APR-2001;
  COURNAG
             MERCK PATENT GmbH (DE)
                      Location/Qualifiers
 DESCURES.
                      1. .1014
      source
                      /organism="Homo sapiens"
                       /mol_type="genomic DNA"
                       /db_xref="taxon:9606"
                       1. .1014
      CDS
                       /note="unnamed protein product"
                       /codon_start=1
                       /protein_id="CAC37735.1"
                       /db_xref="GI:13921509"
                       /translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
                       SLQAMKEILCGTNEKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
                       NILRRQIGTRSPTVEYISAHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
                       FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
                       \stackrel{\cdot }{\text{SENYVTKRQSLKLLGELILDRHNFAIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV}}
                       FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKK
                       TAP"
                                               260 t
                            205 C
                                     209 g
                  340 a
 BASE COUNT
```

Query Match 100.0%; Score 1014; DB 6; Length 1014;
Best Local Similarity 100.0%; Pred. No. 6.9e-238;
Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORICIN

	1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
QУ	1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAAATTTTTTTT
Db	1 ATGAAAAAATGCCTIIGIIIAGIAAATGCAAAAAAAATGCCTTCAGAA 120
Оy	61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC
Db	61 CTGAAAGACAATTIGGCCATTITTT
Qy	121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA
Db	TO GARGE CARCETOTA CAGCAGTGGCCTGCTAGTG 240
Qy Db	181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTAGTG 240 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTAGTG 240
ממ	THE THE STATE OF T
Ωy Db	
,	TARREST AGGREGACTCCTACTGTGGAGTATATTAGT 360
QУ	361 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTTGTGTGTG
Db	A CONTRACTOR ACCCCCACACACTTCCCTTA 420:
Ωy	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCGTAAAGGATATGAAGCCCGTCAAAGGATATGAAGCCCGAAATGCCTTA 420:3
- Ðb _, 	TO A THE THEORY CATED ACCACTTGCCAAAATCATCCTC 480
97 195 :	
$Q_{2'}$	481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
ِم <u>ر</u> ثِ	481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGTGTT
Öλ	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
Db	541 TCAGATGCCTTTGCTACTTTCAAGGATTMC
C.A.	501 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
Dþ	601 TTCTTAGAACAAAATTACGACACTATTTTTGAAAATTACGACGTCAC 720
бĀ	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720:
ďŒ	661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGGTTGATCATGAAC 780
ōλ	721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780.
Dр	721 AACTTTGCCATCATGACAAAGTATTT 840
QY	781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
Db	781 CTCCTTCGGGATAAAAGTCCCAACATCCTTTTTTTTTTT

```
841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
         QУ
      841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
Db
      901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
         Çу
       901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGCACGGATGATGAGCAGTTCGCTGAC 960
Db
       961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
         Qy
       961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Db
RESULT 2
                                            PAT 14-FEB-2001
AR097361
                                DNA
                                      linear
                         1344 bp
        AR097361
LOCUS
        Sequence 2 from patent US 6071721.
DEFINITION
        AR097361
ACCESSION
                GI:12806091
        AR097361.1
VERSION
KEYWORDS '
        Jnknown.
SOURCE
        Unknown.
  ORGANISM
        Unclassified.
          (bases 1 to 1344)
        Tang, Y. Tom., Guegler, K.J., Corley, N.C. and Gorgone, G.A.
REFERENCE
  AUTHORS
         Calcium binding protein
  TITLE
         Patent: US 6071721-A 2 06-JUN-2000;
  COURNAL
              Location/Qualifiers
 VERTURES.
               1. .1344
    source
               /organism="unknown"
                        280 A
                              353 t
                  261 ℃
           450 a
BASE COUNT
 ORIGIN
                                       Length 1344;
                        Score 1010.8; DB 6;
                   99.78;
  Query Match
                        Pred. No. 4.2e-237;
                  99.8%;
  Best Local Similarity
                                                Gaps ..
                                      Indels
                       0; Mismatches
                                    2:
  Matches 1012; Conservative
         1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
          QY
        124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 183
 Db
        Q٧
        dG
        QУ
        Db
        181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
           Qy
        304 CCCCGACAGAACCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 363
 Db
        241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 30.0
           Qy
        364 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423
 Db
```

	301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Qy	301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATTAGT 483 . 424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483 .
Db	THE CHIEF A ACCOUNT TEA AGCCCCACAGATTGCCTTA 420
QУ	361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGIITOOGGATATGIITOOGGATATGIITOOGGATATGIITOOGGATATCCTGTTTATGCTCCTCAAAGGATATGIITOOGGATATGCCTCAAAGGATATGAAGCGCCACAGATTGCCTTA 543 484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCGCCACAGATTGCCTTA 543
Db	THE CANCETT COLOR OF THE COLOR
Όλ	421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGGGAALLACTAGAACCACTTGGGAAATGAACCACTTGCGAAAATCATCCTC 603. 544 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603.
dQ	A CONTROL OF THE CONT
QY	481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGIGGAGITGTGATATTGATATTGCT 663 . 604 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663 .
Db	GEAGAGAGATAAAGTGTTGGTAGCAGAC 600
QУ	541 TCAGATGCCTTTGCTACTTCAAGGATTTACTAACCAGACATAAAGTGTTUUUUUUUUUU
Db	TO BE CAUSED BY A CONTROL OF THE CON
Qу	TTCTTAGAACAAATTACGACACTATTTTTGAAGACTATGAGAAATTOOTTOOTTOOTTOOTTOOTTOOTTOOTTO
Dio	TO THE STATE OF TH
Ωy:	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843.5
745	A COURT A ACTICATION OF THE AC
Qy Dh	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGAGAACCTGATACUTUUUUUUUUUU
Эb	A CCCTTTCATGTTTTAAGGTGTTT 840
QΣ	731 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963 204 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963
Db	841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
57	841 GTGGCCAGTCCTCACHTUTION BY THE STATE OF T
Db	901 CTCATTGACTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
Qy Db	
	961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
QУ	931 GAGAGATIOTTI
Ър	1064 GAGAAGAACTACTIGATTAAACAACTACT
	LT 3 3365
AR20 LOCU	1344 pp
	NITION Sequence 2 from patent US 6365371. SSSION AR203365
VERS	SION AR203365.1 GI:21499736
KEYV	NORDS RCE Unknown

```
Unknown.
 ORGANISM
        Unclassified.
          (bases 1 to 1344)
        Tang, Y. Tom., Guegler, K.J., Corley, N.C. and Gorgone, G.A.
REFERENCE
 AUTHORS
        Calcium binding protein
        Patent: US 6365371-A 2 02-APR-2002;
 TITLE
 JOURNAL
              Location/Qualifiers
FEATURES
              1. .1344
   source
               /organism="unknown"
                              353 t
                        280° g
                 261 C
           450 a
BASE COUNT
ORIGIN
                                       Length 1344;
                                  DB 6;
                        Score 1010.8;
                  99.7%;
 Cuery Match
                        Pred. No. 4.2e-237;
                  99.8%;
 Best Local Similarity
                                                      0;
                                                Gaps
                                              0;
                                       Indels
                                    2:
                          Mismatches
                       0;
            Conservative
 Matches 1012;
         1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
          QУ
      124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183
Db
        QУ
       Dio
       121 CAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAAGAA 180
          181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
           ÇV
        304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 363
 20
        241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
           CA.
        364 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423
 Db
        301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
           ΩУ
        424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483
 DD:
        361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
           QΥ
        484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543
 Db
        421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
           Çy
           CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603
 Db
         481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540....
            Qy-
         504 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663
  Dr.
         541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
            24
         564 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 723
  Db
         601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
  ŌУ
```

	!!!		 TTCAGTCTGAG 783
Db	724 TTC	TTTAGAACAAAATTAGGAAGTGATG	TEGACCGTCAC 720
Qу	661 AAT	TTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCC 	
Db	/84 AA1	CTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAA	CTCATGATGAAC 780.
QУ	721 AAC	CTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAAACO POPUL 	
ďŒ	844 AA	CTTTGCC/TTGTTCATGTT	TTTAAGGTGTTT 840
QУ	781 CT	CCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTT 	 TTTAAGGTGTTT 963
Db	904 CT	CCTTCGGGATAAAAGTCCCAACATGGATGT	AATCACCCCAAA 900
ÕА	343. GT	GGCCAGTCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAA 	AATCAGCCCAAA 300 AATCAGCCCAAA 1023
D.b		THE TAR A CARACCA CCCATGATGAC	CAGTTCGCTGAC 960
QУ	901 CT	CCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGITGIA 	
Dp	1024 C1	AGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGC	CCCTTGA 1014
Qy Db	961 GA 1084 GA	AGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAAAAAAA	
19.7	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		
DEFINITION OF THE PROPERTY OF	SCO ON Hom	10993 1491 bp mRNA linear o sapiens, hypothetical protein FLJ12577, clon	e MGC:15031
. 1	IMA	GE: 3956127, mRNA, complete cus.	ν . ϵ
ACCESSIO		10993 10993.1 GI:15012172	V 4.0
VERSION KUNWORDS	MGC	1	en de la companya de La companya de la co
SOURCE	Hom	no sapiens (human)	4 - A
ORGENI		no sapiens karyota; Metazoa; Chordata; Craniata; Vertebrat	a; Euteleostom1;
,	Man	nmalia; Eutheria; Primates; Catalinini, nome	iae, nomo:
REFERENC	,	(bases 1 to 1491) rausberg,R.	1 × 1 × 1
AUTHOR TITLE	RS Str Dii	rect Submission	ealth. Mammalian
JOURNA	AL Sul	rect Submission bmitted (23-JUL-2001) National Institutes of Ho ne Collection (MGC), Cancer Genomics Office, No stitute, 31 Center Drive, Room 11A03, Bethesda	ational Cancer , MD 20892-2590,
			y
REMAR	•	A H-MGC Project URL: http://mgc.nci.nih.gov ntact: MGC help desk	4×~
COMMENT	∃m	ail: cgapbs-r@mail.nin.gov	
	Ti	ssue Procurement: ATCC	
			um (LLNL)
	CD	ANG HAVE BIOLOGY	,
•	DN	To compending by: Institute for Systems 22	
	ht	NA Sequencing by: Institute for systems biology tp://www.systemsbiology.org ontact: amadan@systemsbiology.org oup Madan, Rachel Dickhoff, Jessica Fahey, Step	

Greene, Mark Ketteman and Anuradha Madan

```
Clone distribution: MGC clone distribution information can be found
        through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
        Series: IRAL Plate: 25 Row: k Column: 12
        This clone was selected for full length sequencing because it
        passed the following selection criteria: matched mRNA gi: 10434146.
               Location/Qualifiers
FEATURES.
               1. .1491
               /organism="Homo sapiens"
   source
               /mol_type="mRNA"
               /db_xref="LocusID:81617"
               /db_xref="taxon:9606"
               /clone="MGC:15031 IMAGE:3956127"
               /tissue_type="Placenta, choriocarcinoma"
               /clone_lib="NIH_MGC_21"
               /lab_host="DH10B-R"
               /note="Vector: pOTB7"
               416. .1285
    CDS
               /codon start=1
               /product="hypothetical protein FLJ12577"
                /protein_id="AAH10993.1"
                /db xref="GI:15012173"
                -translation="MKEILCGTNEKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFE
                EKKDVTQIFNNILRRQIGTRSPTVEYISAHPHILFMLLKGYEAPQIALRCGIMLRECI
                RHEPLAKIILFSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDT
                1FEDYEKLLQSENYVTKRQSLKLLGEL1LDRHNFA1MTKY1SKPENLKLMMNLLRDKS
                PNIQFEAFHVFKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNY
                LIKOIRDLKKTAP"
                                                  Free to sail
                                393 t
                   290 c
                          305 g
                                                  , 1:
            503 a
 MOSS GOLD
 ORIGIN
                                                   1,-16
                                          Length 1491:
                         Score 1009.2; DB 9;
                    99.5%;
                                                   99.7%; Pred. No. 1e-236;
  County Matica
  Best Local Similarity
                                                   Gaps:
                                         Indels
                            Mismatches
                                      3:
  Matches 1011; Conservative
                         0.;
          ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
           272 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 331
 QΥ
         Db
            QУ
 Dr.
         QУ
         Dio
         181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240 -
            452 CCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGGTG 511
  QУ
         241. ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
  Db
            QУ
  Db
         301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
  Qy
```

	531 420
361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	691
632 GCTCATCCTCATATCCTGTTTTTTTTTTTTTTTTTTTTT	480
421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCTAAAAATCATCCTC	751
692 CGTTGTGGGATTATGCTGAGAGTTZ	540
481 TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	811
TO BE A CARA A ACTO TEGERAGO CARA CARA CARA CARA CARA CARA CARA CAR	600, :
	0/1
TESTETT CAACAA AATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	, 600
ACAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
TO THE ATOMICAL CALAGED AND A STREET ATOMICAL CONTRACTOR AND ATOMICA CONTRACTOR AND ATOM	1203: . :
930 BATTATGTTACTAAGAGACAGTGT	C 780
	C 1.004:
TO STRUCK A ACCOUNT CATGITT TAAGGIGIL	T 840
	,1 1111,
TGGGGAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA	11
	AM Tir. 120
GTEAR TTO AGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGC 1GA	AC 300. 1
1172 CTCATTGAGTITCTGAGGATCCGAGACTTGAAGAAAACGGCCCCTTGA 101	4
	5 (1)
1232 GAGAAGAACTACTIGATTAAAAA	
2002 bp DNA TABLE STATE TO STATE THE PRINT OF SYNTHON Primer for synthesizing full-length cDNA and use thereof. SSION BD157871 ION BD157871.1 GI:27863629 ORDS JE 2002191363-A/12714. CE Homo sapiens (human)	JAN-2003
	CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGAAAACGGCCCCTTGA 101 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 101

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
REFERENCE
          Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
 AUTHORS
          Primer for synthesizing full-length cDNA and use thereof
          Patent: JP 2002191363-A 12714 09-JUL-2002;
 TITLE
 JOURNAL
          HELIX RESEARCH INSTITUTE
               Homo sapiens (human)
           OS
COMMENT
               JP 2002191363-A/12714
           PN
               09-JUL-2002
           PD
               28-JUL-2000 JP 2000280990
               TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
           PF
           ΡI
               JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
           PΙ
           PΙ
               KEIICHI NAGAI, TETSUJI OTSUKI
           PΙ
           C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
                C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
            10,
            Primer for synthesizing full-length cDNA and use thereof FH
           PC
                     Location/Qualifiers
                               (127) \cdot (993) \cdot
                CDS
            FT
                    Location/Qualifiers
 FEATURES
                    1. .2002
                    /organism="Homo sapiens"
     BOUTCE
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                                        527
                                463 g
                       418 C
                594 a
 BASE COUNT
 ORIGIN
                                                  Length 2002;
                                Score 992.8; DB 6;
                         97.98;
                                                               .....
                                Pred. No. 1.1e-232;
   Chary Match
                         99.88;
                                                                      0;
                                                               Gaps
   Best Local Similarity
                                                   Indels
                                  Mismatches
                               0;
           994; Conservative
   Matches
            19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAAGACAATTTGGCC 78
              1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60
  \mathbb{C}_F
  DO
            79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
               61 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120
  QУ
           139 CAAGCAA'IGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCCAACAGAAGCAGTG 198
  Tib
               121 CAAGCAATGAAAGAAATTCTGTGTGTGTACAAACGAGAAGAACCCCCAACAGAAGCAGTG 180
  \Omega V
           199 GCTCAGCTAGCACAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
   Do
               181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240
   QУ
            259 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
   Db
                241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300
   QΥ
            319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
   Db.
                QУ
```

	301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360
Db	WAR AND THE COTTEN COTT
QУ	379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
Db	361 TTTATGCTCCTCAAAGGATATGAACCCCCCCCCCCCCCC
QУ	439 AGAGATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498
Db	TO STEER OF THE COURT OF THE CO
Qy	499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
da	A CONCENTRACIONAL CONTRACTOR AND A CONTR
Çy 	TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600
Db .	CHOROLO AND TRATCATATATATATATATATATATATATATATATATATA
QУ	[
D.O .	TOTAL ACCURACY OF THE STATE OF
QY OB	
	THE TABLE AND THE A PROPERTY OF THE PROPERTY O
QY 	AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCGTTIIIIIIIIII
<u> </u>	A CONTROL TIGHT TO A COTTOTT TIGHT TO A COTTOTT TO THE CONTROL THE CONTROL TO THE CONTROL THE CONTROL TO THE CONTROL TO THE CONTROL TO THE CONTROL TO THE CO
Qy	CCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 8493
Db	TO TO THE ACTUAL TERMS OF
Ōλ	259 ACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGTTTTTTTT
ďQ	TO SOLD THE COURT OF THE COURT
Q - y-	919 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 960 901 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 960
ďŒ	
QY QY	979 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Dþ	
RESULT AK0226	linear DRT 01-AUG-2002
LOCUS DEFINI	AR022639 TION Homo sapiens cDNA FLJ12577 fis, clone NT2RM4001047, highly Similar
AECES!	to MO25 PROTEIN
VERSI	ON AK022639.1 GI:10434146
KEYWO SOURC	g yomo sapiens (human)
ORG	ANISM Homo sapiens

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
REFERENCE
            Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,
  AUTHORS
            Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
            Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,
            Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
            Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A.,
            Nakamura, Y., Nagahari, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.
            NEDO human cDNA sequencing project
  TITLE
            Unpublished
  JOURNAL.
            2 (bases 1 to 2002)
REFERENCE
            Isogai, T. and Otsuki, T.
  AUTHORS
             Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
            Direct Submission
  TITLE
             Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
  JOURNAL
             (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
             NEDO human cDNA sequencing project supported by Ministry of
             International Trade and Industry of Japan; cDNA full insert
 COMMEN'T
             sequencing: Research Association for Biotechnology; cDNA library
             construction, 5'- & 3'-end one pass sequencing and clone selection:
             Helix Research Institute (supported by Japan Key Technology Center
             etc.) and Department of Virology, Institute of Medical Science,
             University of Tokyo.
                      Location/Qualifiers
  STATURES
                       1. .2002
                       /organism="Homo sapiens"
      SOURCE.
                       /mol_type="mRNA"
                       /db_xref="taxon:9606"
                       /clone="NT2RM4001047"
                       /cell_line="NT2"
                       /cell_type="teratocarcinoma"
                       /clone_lib="NT2RM4"
                       /note="cloning vector: pME18SFL3~mRNA from uninduced NT2
                       neuronal precursor cells."
                       127. .996
                       /note="unnamed protein product"
       CD3
                       /codon_start=1
                       /protein_id="BAB14147.1"
                       /db_xref="GI:10434147"
                        /translation="MKEILCGTNEKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFE
                       GKKDVTQIFNNILRRQIGTRSPTVEYISAHPHILFMLLKGYEAPQIALRCGIMLRECI
                        \verb"RHEPLVKIILFSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDT"
                        IFEDYEKLLQSENYVTKRQSLKLLGELILDRHNFAIMTKYISKPENLKLMMNLLRDKS
                        PNIQFEAFHVFKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNY
                        LIKQIRDLKKTAP"
                                              527 t
                                     463 g
                            418 C
                   594 a
   BASE COUNT
   ORIGIN
                                                          Length 2002;
                                     Score 992.8; DB 9;
                             97.9%;
     Query Match
                                      Pred. No. 1.1e-232;
                             99.8%;
     Best Local Similarity
                                                                         Gaps
                                                           Indels
                                        Mismatches
     Matches 994; Conservative
                                     0;
              19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAAGACAATTTGGCC 78
                  QУ
                1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60
    Db
```

	TOTAL CANCALGACTE TAAATCACTG 138
δÄ	79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
Db	61 ATTTTGGAAAAGCATG 198
Оλ	61 ATTTTGGAAAGCAAGACAAATTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTO
Dp	121 CAAGCAATGAAAGAATT
Qy	121 CAAGCAATGAAAGAATTCTGTGTGTGTGTGTGACACTGATAGCTGACCTG 258 199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258 191 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240 181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTCTTGAGA 318
Dp	181 GCTCAGCTAGCACAAGAACT
QY	181 GCTCAGCTAGCACAAGAACTCTTGAGA 318 253 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
Do	
QŸ	
· Db.	301 AGACAGATAGGCACTCGGAGTCCTTACGTTGTGGGATTATGCTG 438 379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420
Qy	
ರರ -	361 TTTATGCTCCTCAAAGGATATOATOATOATOATOATCAATCAATCAATCAATCAATC
ĝγ	
2	
QY.	
D1;	GATTTCTTTAAGTACGTGGAGTTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600
Q y	
-D <u>:</u> 5	TTCAAGGATTTACTAACCAGACTAT TTCAAGGATTTACTAACCAGACTAT TTCAAGGATTTACTAACCAGACTAT TTCAAGGATTTACTAAGAGA 678 TTCAAGGATTTACTAAGAGA 678
Ω¥	
Dp	TOTAL CALCADATE AND A CONTROL OF THE
QY	
क्ष	CCCCCACA CCTGAAACTCATGATGAACCTCCTTCGGAACCTTCGAACCTTCGAACCTTCAACCTTCGAACCTTCGAACCTTCGAACACACAACCTTCGAACACAACCTTCGAACACAACCTTCGAACACAACCTTCGAACACAACAACAACAACAACAACAACAACAACAACAAC
Öý-	
Dp.	721 AAGTATATCAGCAAGCCGGACITT 721 AAGTATATCAGCAAGCCGGACITT 722 AAGTATATCAGCAAGCCGACITT 723 CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAA 858
ДŸ	
áŒ	781 CCCAACATCCAGTTTGAAGCCTTTGTTAAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 918 859 ACACAGCCTATTGTGGAGATCCTGTTAAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 900
QУ	859 ACACAGCCTATTGTGGAGATCCTGTTAAAATTATTTTTTTT
Dр	84T ACACAGCCIALLO

```
919 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 978
              901 AGCTTCCANAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 960
QΥ
Db
          979 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
              Qy
          961 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 996
Db
RESULT 7
                                                                 ROD 16-APR-2003
                                                        linear
            Mus musculus RIKEN cDNA 1500031K13 gene, mRNA (cDNA clone MGC:28889
                                                mRNA
 BC016128
 LOCUS
             IMAGE:4911640), complete cds.
 DEFINITION
             BC016128
 ACCESSION
             BC016128.1 GI:16359341
 VERSION
             MGC.
 KEYWORDS
             Mus musculus (house mouse)
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 SOURCE
             Mus musculus
   ORGANISM
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
             Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
  REFERENCE
              Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
    AUTHORS
              Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
              Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
              Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
              Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
              Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
              Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
              Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
              McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
               Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
               Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
               Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S.,
               Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
               Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
               Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
               Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
               Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
               Generation and initial analysis of more than 15,000 full-length
               numan and mouse cDNA sequences
     TITLE
               Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
     OURNAL
                22388257
     MEDLINE
                12477932
      PUBMEL
                   (bases 1 to 1359)
    REFERENCE
                Strausberg, R.
     AUTHORS
                Submitted (22-OCT-2001) National Institutes of Health, Mammalian
                Gene Collection (MGC), Cancer Genomics Office, National Cancer
      TITLE
                Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
      JOURNAL
                NIH-MGC Project URL: http://mgc.nci.nih.gov
      REMARK
                Contact: MGC help desk
                Email: cgapbs-r@mail.nih.gov
    COMMENT
                Tissue Procurement: Jeffrey E. Green, M.D.
                CDNA Library Preparation: Life Technologies, Inc.
```

書きたいと あるか

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 38 Row: m Column: 23 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not dentity to protein. Location/Qualifiers 1. .1359 squirde /organism@"Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="MGC:28889 IMAGE:4911640" /tissue_type="Salivary gland, 10 week old remale mouse" /clone_lib="NCI_CGAP_SG2" /lab host="DH10R" /note="Vector: pCMV-SPORT6" 1. .1359 /gene="1500031K13Rik" /note="synonyms: 4930520C08Rik, 2810425013Rik" /db xref="LocusID:69008" /db_xref="MGI:1916258" 262. .1266 /codon start=1 /product="1500031K13Rik protein" /protein_id="AAH16128.1" /db xref="GI:16359342" /db_xref="LocusID:69008" /translation="MPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQ AMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFNNIL RRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIILFSN QFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQSEN YVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVFKV FVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKKAAP 346 t 294 g 418 a 301 C BASE COUNT 84.9%; Score 860.4; DB 10; Length 1359; Best Local Similarity 90.5%; Pred. No. 3.2e-200; Query March Gaps Indels 96; 0; Mismatches Matches 918; Conservative

FRATURES

ans

ORIGIN.

QУ

1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60

	253 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAAATT 312
Db	TOTAL A COLD CALADA AGA CAGACAAGGCTTCAGAA 120
QУ	61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAAAA
Db	TO A REAL PROPERTY OF THE TRANSPORT OF T
QУ	121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGTG
Db :	THE GRACE GRACE CANCEL GOVERNMENT OF THE STATE OF THE STA
QУ	181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGATTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOT
Db .	433 CCCCCTACAGAAGCAGIGGCICAGCIGGCGGAAAAAAAAAA
QУ	241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
Db	THE GARLEGOGA CTCCGACTCCTACTGTGGAGTATATTAGT 360
Ďλ	201 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTTAGTAGACACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTTAGTAGAGTACATCAGT 612 553 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 612
Db	A CONTRACTOR A CONTRACTOR ACCOCCACAGATTGCCTTA 420
Qy	361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCTTTTTTTT
Db	TARRANGE TECCACATGA ACCACTTGCCAAAATCATCCTC 480
δλ _.	CGTTGTGGGATTATGCTGAGAGAGTGTATTCGACATGAGCGAGATGTATTCGACATGAGCGAGATGTATTCGACATGAGCGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 732 5
Db	GGCTGTGGGATTATGCTAAGAGAGTGTATTGGTAAGAAGAGTGTATTGGT 540
The same	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
Ď.	GGA WITH GTA ACCAGA CATAAAGTGTTGGTAGCAGAC 600
Q¥	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACTTTACTAACCAGACTTTACTAACCAGACTTTACTAACCAGACTTTACTAACCAGACTTTACTTAACCAGACATAAAGTATTGGTAGCAGAC 852 793 TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 852
DL	TA CHARRETTE A A CACTATGAGAAATTGCTTCAGTCTGAG 660
ΩY	601 TTCTTAGAACAAATTACGACACTATTTTTGAAGACTATOMOTA
Db	853 TTCTTAGAACAAAATTATGACACTATTTTTTTTTTTTTT
ΩY	661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
Dρ	TO GENERATE TO A GOAR A GCCCGGAGAACCTGAAACTCATGATGAAC 780
Ο¥	721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGTGTTTTTTTT
Db	973 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAAGCCAGAGAAGCCAGAGAAGCCAGAGAGAGCCAGAGAGAGCCAGAGAGAGCCAGAGAGAGCCAGGAGAGCCAGAGAGCCAGAGAGCCAGAGAGCCAGAGAGCCAGAGAGCCAGAGAGCCAGAGAGCCAGAGAGCCAGAGAGCCAGGAGAGAGCCAGGAGAGCCAGGAGAGCCAGGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGCCAGGAGAGAGAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA
ΰλ	781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
da	1033 CTGCTTCGAGACAAAAGTCCCAACATCGTCACAAAAAATCAGCCCAAA 900
бĀ	841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
Db	

```
Qy
                1153 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGACGACGACGACGAGCAGTTTGCTGAC 1212
Db
                  961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
                          11 MARIA | 11 MARIAN | 11 MARI
QУ
                1213 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1266
Db
RESULT 8
                                                                                                                     ROD 16-APR-2003
                                                                                                     Tinear
 BC016546
                                                                                      mRNA
                                                                  1530 bp
                      Mus musculus RIKEN cDNA 1500031K13 gene, mRNA (cDNA clone MGC:27972
 <u>r</u>ocus
 DEFINITION
                       IMAGE:3595339), complete cds.
                       BC016546
 ACCESSION
                       BC016546.1 GI:16741456
 VERSION
  KEYWORDS
                       MGC.
                       Mus musculus (house mouse)
  SOURCE
                       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                       Mus musculus
     CRGANISM
                       Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                             (bases 1 to 1530)
                       Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
  PEFERENCE
                        Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
      AUTHORS
                        Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
                        Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
                        Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
                        Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
                        Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
                         Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.
                         Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
                         McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
                         Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
                         Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, K., Gibbs, R.A.,
                         Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S.,
                         Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
                         Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
                         Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
                          Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
                          Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
                          Generation and initial analysis of more than 15,000 full-length
        TITLE
                          human and mouse cDNA sequences
                          Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
        JOURNAL
                          22388257
        MEDLINE
                          12477932
          PUBMED
                           2 (bases 1 to 1530)
     REFERENCE
                           Strausberg, R.
         AUTHORS
                           Submitted (31-OCT-2001) National Institutes of Health, Mammalian
         TITLE
                           Gene Collection (MGC), Cancer Genomics Office, National Cancer
         JCURNAL
                           Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
                           NIH-MGC Project URL: http://mgc.nci.nih.gov
         REMARK
                           Contact: MGC help desk
      COMMENT
                           Email: cgapbs-r@mail.nih.gov
```

Tissue Procurement: Jeffrey Green M.D.

```
CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNR Sequencing by: Baylor College of Medicine Human Genome
          Sequencing Center
          Center code: BCM-HGSC
          .Web site: http://www.hgsc.bcm.tmc.edu/cdna/
          Contact: amg@bcm.tmc.edu
           Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
           Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
           A.N., Gibbs, R.A.
           Clone distribution: MGC clone distribution information can be found
           through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
           Series: IRAK Plate: 35 Row: m Column: 15
           This clone was selected for full length sequencing because it
           passed the following selection criteria: Similarity but not
           identity to protein.
                    Location/Qualifiers
FEATURES
                    1. .1530
    source.
                    /organism="Mus musculus"
                    /mol type="mRNA"
                    /strain="FVB/N"
                    /db_xref="taxon:10090"
                    /clone="MGC:27972 IMAGE:3595339"
                     /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
                     ductal carcinoma. 5 month old virgin mouse."
                     /clone_lib="NCI_CGAP_Mam6"
                     /lab host="DH10B"
                     /note="Vector: pCMV-SPORT6".
                     1. .1530
     sene
                     /gene="15\0031K13Rik"
                     /note="synonyms: 4930520C08Rik, 2810425013Rik"
                     /db_xref="LocusID:69008"
                     /db_xref="MGI:1916258"
                     279. .1283
                     /codon_start=1
                     /product="1500031K13Rik protein"
                     /protein_id="AAH16546.1"
                     /db_xref="GI:16741457"
                     /db_xref="LocusID:69008"
                     -translation="MPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQ
                     AMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFNNIL
                      RRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIILFSN
                      QFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQSEN
                      YVTKRQSLKLRGELII.DRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVFKV
                      FVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKKAAP
                                            393 t
                                   326 g
                          313 C
                 498 a
 DASE- COUNT
 ORIGIN
                                  Score 858.8; DB 10; Length 1530;
                           84.7%;
   Query Match
                                   Pred. No. 8e-200;
                           90.4%;
   Best Legal Similarity
                                                                     Gaps .
                                                        Indels
                                  0; Mismatches
                                                   97;
   Matches 917; Conservative
             1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
               270 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 329
```

ÒУ

DD

Ōλ		CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGGCTTCAGAA	
Db		CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC	
Qу	•	GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	
Db	,	CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGGTAGTG	
Qy Db		CCCCCACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCTGTGGTGTGTGT	
Qу	241	ACACTGATAGCTGACCTGCAGCTGATAGACT'ITGAGGGAAAAAAAGATGTGACCCAGATA	300:
Db	- 510	ACACTGATAGCTGATGTGATGTGATGTGACCTGATGACCTGATGACCTGATGACCTGATGACCTGATGACCTGATGACCTGATGACCTGATGAGGGAAAAAAAGATGTGACCCAGATA	569:
Qу	301	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	360
Db	570	TTCAACAACATCC'IGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT	629
Q'Y	361	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	420:
áT	030	TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA	689-
Çу	,21	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	480.
n.j.,	390		74: 9
27	481		540
DÃ.	25 0	TITICIAAT CAGACITICAGACITICAGACITICAGACITICAGACACTATGATATCGCT TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT	809
୍ଦି ପୁଧ	541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	600:
DD	ខ±៌0		869
Qy	5.01	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	56.0
DЬ	870		929
Qý	561	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	72.0
Db.	930		989
ÇУ	721	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	780∢
Db.	990	AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC	1049
Qγ	783	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	840.
Db	1050		1109
Qу	84.	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	. 900°.
Db	1110		. 1169

÷

1<u>1</u>

..

```
901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
Qу
             1170 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGACGACGACCAGCAGCAGTTTGCTGAC 1229
Db
         961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Qy
           1230 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1283
Db
RESULT 9
BD147463
                                                            PAT 17-JAN-2003
                                            DNA
                                                    linear
                                   822 bp
           BD147463
LOCUS
           Primer for synthesizing full-length cDNA and use thereof.
DEFINITION
ACCESSION
           BD147463
           BD147463.1 GI:27853221
VERSION
           JP 2002191363-A/2306.
KEYWORDS
           Homo sapiens (human)
SOURCE
  ORGANISM
           Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           1 (bases 1 to 822)
REFERENCE
           Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
  AUTHORS
           Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
           Primer for synthesizing full-length cDNA and use thereof
  TITLE
           Patent: JP 2002191363-A 2306 09-JUL-2002;
  JOURNAL
           HELTK RESEARCH INSTITUTE
               Homo sapiens (human)
COMMENC
                JP 2002191363-A/2306
            M
                09-JUL-2002
            PD
                28-JUL-2000 JP 2000280990
           \mathbb{P}^{V}
                TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
           DI.
            ĎΙ.
               SAITO,
                JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
            ÞΤ
            PI KEIICHI NAGAI, TETSUJI OTSUKI
            C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
            PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH
                      Location/Qualifiers
            FT
                 source
                                /organism='Homo sapiens (human)'.
            \overline{a}
                    Location/Qualifiers
 FEATURES
                     1. .822
     source
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                                                    3 others
                                         216 t
                                 171 g
                268 a
 BASE COUNT
 ORIGIN
                          76.0%; Score 770.6; DB 6; Length 822;
   Query Match
                         98.5%; Pred. No. 3.4e-178;
   Best Local Similarity
                                                10; Indels
                                                                  Gaps 2;
                                0; Mismatches
   Matches 798; Conservative
           19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAAGACAATTTGGCC 78
 Qy:
              1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60
 Db
```

Ωy	79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138	
Db	61 ATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAAAAGACAGAC	
Qy	139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAGAACCCCCAACAGAAGCAGTG 198 	
Db	258	
Qy Db	199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCTGGTGACACTAGCTGACCTG 240 181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240	
Qy	259 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318	
Dp	241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATTTTTTTT	
Qу	319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378	
ΔŒ	TO THE TIPE A GCCCCACACACTTCCCTTACGTTGTGGGATTATGCTG 438	
DP ÖÄ.	361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420	
. <u>O</u> Z	AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498	
	GATTTCTTTAAGTACGTGGAGITGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558-4-4-4-4-4-4-4-4-4-4-4-4-4-4-4-4-4-4-	
Qy'	TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 61855	
Qγ	619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678	
ΩУ	579 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738	
Db.	GG1 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720	
O.V.	AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798	:
Db Qy	799 CCCAACATCCAGTTTGAAGCCTTTCATGTT 828	:
Db	7.75 000.110.22 4.2-1-1	

RESULT 10 BD079551 LOCUS

```
DEFINITION Cancer-associated nucleic acids and polypeptides.
          BD079551
ACCESSION
          BD079551.1 GI:22625154
VERSION
          JP 2001516009-A/217.
KEYWORDS
          Homo sapiens (human)
SOURCE
          Homo sapiens
 ORGANISM
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 831)
REFERENCE
          Old, L.J., Scanlan, M.J., Stockert, E., Gure, A., Chen, Y.T., Gout, I.,
 AUTHORS
          OGhare, M., Obata, Y., Pfreundschuh, M., Tureci, O. and Sahin, U.
          Cancer-associated nucleic acids and polypeptides
 TITLE :
          Patent: JP 2001516009-A 217 25-SEP-2001;
  JOURNAL
          LUDWIG INSTITUTE FOR CANCER RESEARCH
              Homo sapiens (human)
COMMENT
              JP 2001516009-A/217
          ΡN
              25-SEP-2001
              15-JUL-1998 JP 2000503425
          PF
                                                     60/061599 PR
                             08/896164,10-OCT-1997 US
              17-JUL-1997 US
          PR
                         60/061765,10-OCT-1997 US
                                                 08/948705 PR
          10-OCT-1997 US
                                                              LLOYD
                                                 09/102322 PI
                         9721697.2,22-JUN-1998 US
          11-OCT-1997 GB
          J OLD, MATTHEW J SCANLAN, ELISABETH STOCKERT, ALI GURE, YAO FI
                                                               TSENG
          CHEN.
              IVAN GOUT, MICHAEL O'HARE, YUICHI OBATA, MICHAEL PFREUNDSCHUH, PI
            OZLEM TURECI,
              UGUR SAHIN
          PI
          PC
          001N33/574,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61K48/00, PC 3
           A61P35/00,
          FC C07K14/82,C07K16/32,C12N15/09//C07K16/46,C12P21/08,A61K37/02,
               C12N15/00
           ЪС
              Cancer-associated nucleic acids and polypeptides. FH Key
                 Location/Qualifiers
                            1. .831
          TT
               source
                            /organism='Homo sapiens (human)'.
           T
                  Location/Qualifiers
THRIURES
                  1. .831
     source
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db xref="taxon:9606"
                                     209 t
                                              5 others
                             167 g
                      165 c
BASE COUNT
              285 a
 CRIGIN
                                               Length 831;
                             Score 684.6; DB 6;
                      67.5%;
  Query Match
                      96.1%; Pred. No. 4.1e-157;
  Best Local Similarity
                            0; Mismatches
                                                           Gaps
                                               Indels
         764; Conservative
           1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
 QУ
             37 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 96
 ď
          ÇY
             Db
         ÇУ
```

Db		157	GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	216
Çу			CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG	
Db			CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTAGTG	
Qу		241	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA	300
Db		277	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA	336
Qу		301	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	360
Db		337	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	396
QУ		361	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	420
Db			GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCCACAGATTGCCTTA	
	_		CONTROL CONTRATOR CONTRATOR CARGA A TGTATT CGA CATGA A CCA C'I TGC CA A A A TCAT CCT C	
ÇУ	•		CGTTGTGGGATTATGCTGAGAGATTTTTTTTTTTTTTTT	
Db	· · ·			
Qγ			TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	
DF)		TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTIGATATIGCT	
Qy	,	543	. 'rCAGATGCCTTTGCTACTTTCAA-GGATTTACTAACCAGACATAAAGTGTTGGTAGC-AG	598; ;
DE)	5° 7	CAGATGCCTTTGCTACTTTCAAGGGATTTACTAACCAGACATAAAGTGTTGGTAGCAAG	636 <i>i i</i>
G _i g		.5. <u>@</u> 6	ACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTG	65 8 %
Dì	· >	635	ACTICITAGAACAAAATTACGACACTANTTTTGAAGACTATGAAAATTGCTTCAGTCTG ACTTCTTAGAACAAAATTACGACACTANTTTTGAAGACTATGAAAATTGCTTCAGTCTG	596 <i>-</i>
Q ₂	5	653	AG-AATTATGTTACTAAGAGACAGTCTTTAAAG-CTGCTAGGGGAGCTGATCCTGGACCG	716
D!		69	AG-AATTATGITACTAAAAATTATTATAAAATTATTATAAAAATTATGITACTAAAAAAATTATGITACCAAGAGACAGTCCTTAAAGCCTGCTAAGGGAACTGATTCTGGACCG	756
Q	,		7. TCACAACTTTGCCATC-ATGACAAAG'IATATCAGCAAGCCGGAGAACCTGAAACTCA	
Dì				
			3 TGATGAACCTCCTTC 787	
Q)	ž.			•• • .
D	0	81	7 GGAGGAACCTCCTTC 831	
A L D	ESULT 11 X061831 OCUS EFINITIO	ON	AX061831 1026 bp DNA linear PAT 24-JAN Sequence 1 from Patent WO0078947. AX061831	I-2001
ACCESSION VERSION			AX061831.1 GI:12539911	
	EYWORDS OURCE		Homo sapiens (human)	
	ORGANIS	SM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleosto	omi;

```
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
          den Daas, I., Fischer, V., Seyfried, C. and von Melchner, L.
 AUTHORS
          Head trauma induced cytoplasmatic calcium binding protein
 TITLE
          Patent: WO 0078947-A 1.28-DEC-2000;
 JOURNAL
          MERCK PATENT GmbH (DE)
                  Location/Qualifiers
FEATURES
                  1. .1026
    source
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db xref="taxon:9606"
                  1. .1026
    നാട
                  /note="unnamed protein product"
                  /codon start=1
                  /protein_id="CAC25030.1"
                   /db xref="GI:12539912"
                   /translation="MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVS
                  KNLVAMKEILYGTNEKEPQTEAVAQLAQELYNSGLLSTLVADLQLIDFEGKKDVAQIF
                  NNILRRQIGTRTPTVEYICTQQNILFMLLKGYESPEIALNCGIMLRECIRHEPLAKII
                  LWSEQFYDFFRYVEMSTFDIASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYEKLL
                  HSENYVTKRQSLKLLGELLLDRHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFH
                  VFKVFVANPNKTQPILDILLKNQAKLIEFLSKFQNDRTEDEQFNDEKTYLVKQIRDLK
                   RPAOOEA"
                                      265 t
                      199 c
                              203 g
              359 a
BASE COUNT
ORIGIN -
                                                Length 1026;
                              Score 582.6; DB 6;
                       57.5%;
  Query Match
                              Préd. No. 4.3e-132;
                                                             ...
  West Local Similarity.
                       74.78;
                                                                    1; -:
                             0; Mismatches 244;
                                                             'Gaps.
                                                 Indels
         748; Conservative
  colochés
          18 GITTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
Qy
             1111
          12 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGCATGGC 71
                                     -AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
          78 CATTTTGGAAAAGCAAGAC-
Qy
                                            1111 11111 111111111111111
                                      72 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 131
ďQ
         Ç.,
                        | \cdot | \cdot |
         132 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 191
 Db
          389 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
 Qу
             192 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 251
CC.
          249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
                   Qу
          252 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 311
 Db
          309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
 () y
                                   31.2 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCAGCCAACA 371
 DЬ
          369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
 Ωy
                    1111 111 1 1 11111 11111 1 111 1 11
                                                         - 11
          372 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 431
```

Db

Qy Db		GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 	
QУ	439	TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC	548
Db		CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA	
Qу		CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGAGAGATTTTTGGA	
Qy Db		ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT	
Qv.		9 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTŢGC 	
Dp CA Dp	72	9 CATCATGACAAAGTATATCAGCAAG&CGGAGAACCTGAAACTCATGATGAACCTCCTTCG 	788 ·
02° Db	- 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1	9 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAG 	85 1 92
∑. ∑. Db		* TCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 	
97 97 20		9 GTTTC'IGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 	
ďG		59 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009 	
AMO LOC DEF ACC VER KEY SOU	TNITION	AX082322 3281 bp DNA linear PAT 28-FEB Sequence 26 from Patent WO0111032. AX082322 AX082322.1 GI:13184499 Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleoste Nammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	FERENCE AUTHORS	1 Hodgson,D.M., Lincoln,S.E., Russo,F.D., Spiro,P.A., Banville, Bratcher,S.R., Dufour,G.E., Cohen,H.J., Rosen,B.H., Chalup,M. Hillman,J.L., Jones,A.L., Yu,J.Y., Greenawalt,L.B., Panzer,S.	U · /

```
Roseberry, A.M., Wright, R.J. and Daniels, S.E.
                   Secretory molecules
  TITLE
                   Patent: WO 0111032-A 26 15-FEB-2001;
   JOURNAL
                    Incyte Genomics, Inc. (US)
                                  Location/Qualifiers
FEATURES
                                  1. .3281
        source
                                   /organism="Homo sapiens"
                                   /mol_type="genomic DNA"
                                   /db xref="taxon:9606"
                                   /note="Incyte ID'No: 481257.3"
                                                                       990 t
                                         601 C
                                                        676 q
                         1014 a
BASE COUNT
ORIGIN
                                                                                           Length 3281;
                                                        Score 582.6; DB 6;
                                           57.5%;
   Query Match
                                                        Pred. No. 4.3e-132;
                                           74.7%;
   Best Local Similarity
                                                                                                                                1;
                                                                                           Indels
                                                                                                            9;
                                                                                                                  Gaps
                                                             Mismatches 244;
                                                       0;
                 748; Conservative
                  18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
QУ
                        101 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 160
Dh
                  78 CATTTTGGAAAAGCAAGAC-----AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
 QУ
                                                                                  1111 11111 1111111111111111
                                                                       11111
                           161 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAAGGCTACAGAAGAAGTTTC 220
 7) }
                 THE ELECTRONICATION OF THE FIRST STATE OF THE STATE OF TH
 QУ
                 CAAAAATCTGGTTGCCATGAAAGAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 280
 209 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
                         ÉIRIGH A BHA TA TR'ATHAILIGE A' ANN A AN
                                                                                                              AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 340
 Db
                  308. AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308.
 QУ
                                   341 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 400
 Db
                  309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
 QУ
                                401 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 460
 Db
                  369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
                                      ÇУ
                  461 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 520
 Db
                  429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
  Qy
                                                               521 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 580
  Db
                  439 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
  Çγ
                                                                              \Pi
                  581 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 640
  DD
                   549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
  Qу
                                                                                                      11111 11 11
                           641 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA 700
```

Db

Qy	609	ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTG	AGAATTATGT 6 GAAAATTATGT 7	760
Dio	701	ACAGC'ATTATGATAGATTTT'CAGTGAATATGAGAAGTM022		700
ОÃ	669	TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTC	CACAACTTTGC CACAACTTCAC	820
Db	7.61			700
QУ	729	CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGA	AACCTCCTTCG	880.
Db	821			0.4.0
QY .	789	GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTG	TTTGTGGCCAG	940
Db	881	AGACAAAAG I COCAAGATT OO TOO TOO TOO TOO TOO TOO TOO TOO T	AAACTCATTGA	908
Qy Do	849	TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCC	AAACTCATAGA	1000
20		GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCT	GACGAGAAGAA	968
Qy Dib	909 1000	GTTTCTGAGCAGCTTCCAARTOTATATATATATATATATATATATATATATATATAT	[1060
		9 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009	+	. 6
Qy		CTACTIGATIAGATCAGATCAGGGATTTGAAGAGACCAGCTC 1101		1040 0
Db	500 	(MATTAGEMENT)		of a
			867500 (177)	. *
	13.			
7 (1952) reside (N. 2	: 76	nua linoar	DDT 22÷JAN	
rebido.	570	3761 bp mRNA linear slope MGC:21631 IMAGE:	PRI 22-JAN	-2002
DEFIN		SC920570 Homo sapiens, MO25 protein, clone MGC:21631 IMAGE:	PRI 22-JAN	-2002
DUFIN	NOLTI	Homo sapiens, MO25 protein, clone MGC:21631 IMAGE: complete cds.	PRI 22-JAN	-2002
DEFIN: ACCES	NOLTI NOIS	Homo sapiens, MO25 protein, clone MGC:21631 IMAGE: complete cds. BC020570	PRI 22-JAN	-2002
DUFIN ACCES VERSI	NOITI	### SC020570 Homo sapiens, MO25 protein, clone MGC:21631 IMAGE: complete cds. ###################################	PRI 22-JAN	-2002
DEFIN ACCES VERSIO KRYWO	TTION SION ON '	Homo sapiens, MO25 protein, clone MGC:21631 IMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC.	PRI 22-JAN	-2002
DEFINIACCES: VERSION KEYWOO SOURCE	ITION SION ON RDS E	Homo sapiens, MO25 protein, clone MGC:21631 IMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human)	PRI 22-JAN 4397573, mRN	-2002 A,
DEFINIACCES: VERSION KEYWOO SOURCE	TTION SION ON '	Homo sapiens, MO25 protein, clone MGC:21631 TMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida	PRI 22-JAN 4397573, mRN 43 PRI	-2002 A,
DEFINI ACCES VERSIO KEYWO SOURC CRG	ITION SION ON RDS E ANTSM	Homo sapiens, MO25 protein, clone MGC:21631 TMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761)	PRI 22-JAN 4397573, mRN 43 PRI	-2002 A,
DEFINIACCES VERSION KEYWOO CORG	SION ON RDS E ANTSM	Homo sapiens, MO25 protein, clone MGC:21631 TMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R.	PRI 22-JAN 4397573, mRN a; Euteleosto ae; Homo	-2002 A, omi;
DEFINIACCES: VERSION KEYWOO CORG	ITION SION ON RDS E ANTSM EENCE	Homo sapiens, MO25 protein, clone MGC:21631 IMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R. Direct Submission	PRI 22-JAN 4397573, mRN a; Euteleosto ae; Homo	-2002 A, omi;
DEFINIACCES VERSION KEYWO SOURC ORG REFER	SION ON RDS E ANTSM EENCE HORS	Homo sapiens, MO25 protein, clone MGC:21631 IMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R. Direct Submission Submitted (03-JAN-2002) National Institutes of Hea	PRI 22-JAN 4397573, mRN a; Euteleosto ae; Homo	-2002 A, omi; ian r
DEFINITION ACCESS VERSION SOURCE CRG REFER AUT JOU	ITION SION ON RDS E ANTSM EENCE	Homo sapiens, MO25 protein, clone MGC:21631 IMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R. Direct Submission Submitted (03-JAN-2002) National Institutes of Hea	PRI 22-JAN 4397573, mRN a; Euteleosto ae; Homo	-2002 A, omi; ian r
DEFINIACCES VERSION KEYWO SOURC ORG REFER	SION ON RDS E ANTSM EENCE HORS	Homo sapiens, MO25 protein, clone MGC:21631 TMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg,R. Direct Submission Submitted (03-JAN-2002) National Institutes of Headen Collection (MGC), Cancer Genomics Office, National Institute, 31 Center Drive, Room 11A03, Bethesda,	PRI 22-JAN 4397573, mRN a; Euteleosto ae; Homo	-2002 A, omi; ian r
DEFINIACCES: VERSION SOURCES ORG REFER AUT JOS	ITION SION ON RDS E ANTSM EENCE HORS TLE JRMAL	Homo sapiens, MO25 protein, clone MGC:21631 TMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R. Direct Submission Submitted (03-JAN-2002) National Institutes of Headen Collection (MGC), Cancer Genomics Office, National Institute, 31 Center Drive, Room 11A03, Bethesda, USA NIH-MGC Project URL: http://mgc.nci.nih.gov	PRI 22-JAN 4397573, mRN a; Euteleosto ae; Homo	-2002 A, omi; ian r
DEFINIACCES: VERSION REFERENCES REFERENCES AUTOMORE REFERENCES REF	ITION SION ON RDS E ANTSM EENCE HORS FLE IRMAL	Homo sapiens, MO25 protein, clone MGC:21631 TMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R. Direct Submission Submitted (03-JAN-2002) National Institutes of Headen Collection (MGC), Cancer Genomics Office, National Institute, 31 Center Drive, Room 11A03, Bethesda, USA NIH-MGC Project URL: http://mgc.nci.nih.gov	PRI 22-JAN 4397573, mRN a; Euteleosto ae; Homo	-2002 A, omi; ian r
DEFINIACCES: VERSION SOURCES ORG REFER AUT JOS	ITION SION ON RDS E ANTSM EENCE HORS FLE IRMAL	Homo sapiens, MO25 protein, clone MGC:21631 IMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R. Direct Submission Submitted (03-JAN-2002) National Institutes of Headen Collection (MGC), Cancer Genomics Office, National Institute, 31 Center Drive, Room 11A03, Bethesda, USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-r@mail.nih.gov	PRI 22-JAN 4397573, mRN a; Euteleosto ae; Homo	-2002 A, omi; ian r
DEFINIACCES: VERSION REFERENCES REFERENCES AUTOMORE REFERENCES REF	ITION SION ON RDS E ANTSM EENCE HORS FLE IRMAL	Homo sapiens, MO25 protein, clone MGC:21631 TMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R. Direct Submission Submitted (03-JAN-2002) National Institutes of Headen Collection (MGC), Cancer Genomics Office, National Institute, 31 Center Drive, Room 11A03, Bethesda, USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-r@mail.nih.gov	PRI 22-JAN 4397573, mRN 4397573, mRN a; Euteleosto ae; Homo alth, Mammal tional Cance MD 20892-25	-2002 A, omi; ian r
DEFINIACCES: VERSION REFERENCES REFERENCES AUTOMORE REFERENCES REF	ITION SION ON RDS E ANTSM EENCE HORS FLE IRMAL	Homo sapiens, MO25 protein, clone MGC:21631 TMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R. Direct Submission Submitted (03-JAN-2002) National Institutes of Headen Collection (MGC), Cancer Genomics Office, National Institute, 31 Center Drive, Room 11A03, Bethesda, USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC	PRI 22-JAN 4397573, mRN a; Euteleosto ae; Homo. alth, Mammal tional Cance MD 20892-25	-2002 A, omi; ian r 90,
DEFINIACCES: VERSION REFERENCES REFERENCES AUTOMORE REFERENCES REF	ITION SION ON RDS E ANTSM EENCE HORS FLE IRMAL	Homo sapiens, MO25 protein, clone MGC:21631 IMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R. Direct Submission Submitted (03-JAN-2002) National Institutes of Head Gene Collection (MGC), Cancer Genomics Office, Nat Institute, 31 Center Drive, Room 11A03, Bethesda, USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortiu	PRI 22-JAN 4397573, mRN 4397573, mRN a; Euteleosto ae; Homo. alth, Mammal tional Cance MD 20892-25	-2002 A, ian r 90,
DEFINIACCES: VERSION REFERENCES REFERENCES AUTOMORE REFERENCES REF	ITION SION ON RDS E ANTSM EENCE HORS FLE IRMAL	Homo sapiens, MO25 protein, clone MGC:21631 IMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R. Direct Submission Submitted (03-JAN-2002) National Institutes of Head Gene Collection (MGC), Cancer Genomics Office, Nat Institute, 31 Center Drive, Room 11A03, Bethesda, USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortiu	PRI 22-JAN 4397573, mRN 4397573, mRN a; Euteleosto ae; Homo. alth, Mammal tional Cance MD 20892-25	-2002 A, ian r 90,
DEFINIACCES: VERSION REFERENCES REFERENCES AUTOMORE REFERENCES REF	ITION SION ON RDS E ANTSM EENCE HORS FLE IRMAL	Homo sapiens, MO2!5 protein, clone MGC:21631 IMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R. Direct Submission Submitted (03-JAN-2002) National Institutes of Headene Collection (MGC), Cancer Genomics Office, Nat Gene Collection (MGC), Cancer Genomics Office, Nat Institute, 31 Center Drive, Room 11A03, Bethesda, USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortiu DNA Sequencing by: Sequencing Group at the Stanfor Center, Stanford University School of Medicine, Secuency	PRI 22-JAN 4397573, mRN 4397573, mRN a; Euteleosto ae; Homo. alth, Mammal tional Cance MD 20892-25	-2002 A, ian r 90,
DEFINIACCES: VERSION REFERENCES REFERENCES AUTOMORE REFERENCES REF	ITION SION ON RDS E ANTSM EENCE HORS FLE IRMAL	Homo sapiens, MO25 protein, clone MGC:21631 IMAGE: complete cds. BC020570 BC020570.1 GI:18088260 MGC. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominida 1 (bases 1 to 3761) Strausberg, R. Direct Submission Submitted (03-JAN-2002) National Institutes of Head Gene Collection (MGC), Cancer Genomics Office, Nat Institute, 31 Center Drive, Room 11A03, Bethesda, USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortiu	PRI 22-JAN 4397573, mRN 4397573, mRN a; Euteleosto ae; Homo. alth, Mammal tional Cance MD 20892-25	-2002 A, ian r 90,

```
Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
          P. M.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 27 Row: d Column: 16
          This clone was selected for full length sequencing because it
          passed the following selection criteria: Hexamer frequency ORF
          analysis, GenomeScan gene prediction.
                   Location/Qualifiers
FEATURES
                   1. .3761
    cource
                   organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="LocusID:51719"
                   /db xref="taxon:9606"
                   /clone="MGC:21631 IMAGE:4397573"
                   tissue type="Duodenum, adenocarcinoma"
                   /clone_lib="NIH_MGC_88" .
                   /lab_host="DH10B"
                   /note="Vector: pCMV-SPORT6"
                   325. .1350
    CDS
                   /codon_start=1
                   /product="MO25 protein"
                   /protein_id="AAH20570.1"
                   /db xref="GI:18088261"
                   /translation="MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVS
                   KNLVAMKEILYGTNEKEPOTEAVAQLAQELYNSGLLSTLVADLQLIDFEGKKDVAQIF
                   NNILRRQIGTRTPTVEYICTQQNILFMLLKGYESPEIALNCGIMLRECIRHEPLAKII
                   LWSEQFYDFFRYVEMSTFDIASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYEKLL
                   HSENYVTKRQSLKLLGELLLDRHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFH
                   VFKVFVANPNKTQPILDILLKNQAKLIEFLSKFQNDRTEDEQFNDEKTYLVKQIRDLK
                   RPACOEA"
                                       1073 t
                       709 C
                               808 g
SABE COUNT
              1171 a
ORIGIN
                                                  Length 3761;
                               Score 582.6; DB 9;
                        57.5%;
  Query Match
                        74.7%; Pred. No. 4.3e-132;
  Best Local Similarity
                              0; Mismatches 244;
                                                   Indels
  Matches 748; Conservative
          18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAAGACAATTTGGC 77
Qy
             336 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGCATGGC 395
Db
                                      -AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
          78 CATTTTGGAÄAAGCAAGAC--
Qy
                                            396 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAAGGCTACAGAAGAAGTTTC 455
Dia
         QУ
                         456 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 515
DO
          189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
 O_{\Sigma}
              1111111111 18111 11 11 1818111111 1 11111 11
          516 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 575
 Db
          249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
 QУ
```

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Db		
QУ	309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368	
Db		
GХ	369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428	
Db	696 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 755	
Qу	429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488	
Db	756 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 815	
Qу	489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548	
Db -		
СХ	549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608	-
Db	GALLA TERA CONCINCTATTTTTCA AGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668	
DP GA	609 ACAAAATTACGACACTATTTTTGAAGACTATTTTTGAAGACTATTTTTGAGAGATTATGATTAGTTTTTTGAGTGAATATGAGAAGTTACTTCAGTAGAAAATTATGT 995	1
Q y	TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728	Ç
:Dh	GACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGA	
∰.	CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 783	
Db.	1056 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACGTGGTGGG	
Ò.	789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAG 848	
Dh	·	
őň	849 TCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908 	
מת		
Qy Db	909 GTTTCTGAGCAGCTTCCAAAAAGAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968 	
Qy ·	959 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009	,
Db		٠.

RESULT 14

AF151824

linear

PRI 18-MAY-2000

LOCUS AF151824 1680 bp mRNA lines
DEFINITION Homo sapiens CGI-66 protein mRNA, complete cds. AF151824

ACCESSION

```
AF151824.1 GI:4929600
VERSION
KEYWORDS
           Homo sapiens (human)
SOURCE
           Homo sapiens
 ORGANISM
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              (bases 1 to 1680)
REFERENCE
           Lai, C.H., Chou, C.Y., Ch'ang, L.Y., Liu, C.S. and Lin, W.
  AUTHORS
           Identification of novel human genes evolutionarily conserved in
  TITLE
           Caenorhabditis elegans by comparative proteomics
           Genome Res. 10 (5), 703-713 (2000)
  JOURNAL
           20272150
  MEDLINE
           10810093
  PUBMED
           2 (bases 1 to 1680)
REFERENCE
           Lin, W.-C.
  AUTHORS
           Direct Submission
  TITLE
           Submitted (17-MAY-1999) Institute of Biomedical Sciences, Academia
  JOURNAL
           Sinica, No. 128, Sec. II, Academia Road, Taipei 115, Taiwan
                   Location/Qualifiers
FEATURES
                   1. .1680
    source
                    /organism="Homo sapiens"
                    /mol type="mRNA"
                    /db xref="taxon:9606"
                    1. .1026
     CDS
                    /codon start=1
                    /product="CGI-66 protein"
                    /protein_id="AAD34061.1"
                    /db xref="GI:4929601"
                    /translation="MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVS
                    KNLVAMKEILYGTNEKEPQTEAVAQLAQELYNSGLLSTLVADLQLIDFEGKKDVAQIF
                    NNILRRQIGTRTPTVEYICTQQNILFMLKGYESPEIALNCGIMLRECIRHEPLAKII
                    LWSEQFYDFFRYVEMSTFDIASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYEKLL
                    HSENYVTKRQSLKLLGELLLDRHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFH
                    VFKVFVANPNKTQPILDILLKNQAKLIEFLSKFQNDRTEDEQFNDEKTYLVKQIPDLK
                    RPACOEA"
                                358 q 458 t
                       324 C
               540 a
BASE COUNT
ORIGIN
                                                   Length 1680;
                                Score 581.6; DB 9;
                         57.4%;
  Query Match
                        74.7%; Pred. No. 7.5e-132;
  Best Local Similarity
                                                                :Gaps
                                                    Indels
                               0; Mismatches 244;
  Matches 747; Conservative
           19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAAGACAATTTGGCC 78
 Qу
             13 TTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGCT 72
 Db
                                      -AAAAAGACAGACAAGGCTTCAGAAGAAGTGTCT 129
           79 ATTTTGGAAAAGCAAGAC---
 QY
                                       4444 - 111 - 1111 - 1141 - 1141 - 11
           73 GTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAAGGCTACAGAAGAAGTTTCC 132
 Db
          \Sigma \lambda
                         133 AAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGACA 192
 Db
          190 GAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATA 249
 ÇУ
            193 GAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGTA 252
 Db
```

ДУ		GCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAAC	309
)b		GCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAAGACGTGGCTCAAATTTTCAACAAT	
QУ .		ATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCT	
Db		ATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACAG	
Qy		CATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGG	
αíΩ		NATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGAA	
Qy		ATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAAT	
D)o			
QΫ́		CAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCC	
Db			
Qy		TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAA	
Dp.			
Q5 * Vol. 1 0 1 1		CAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTT	
7.9-		ACTAAGAGACAGTCTTTAAAGCTGCTACGGGAGCTGATCCTGGACCGTCACAACTTTGCC	
ිලිලි වර්		ACTANGAGACAGTCTTTAAAGCTGCTACGGAGCTACTACTAGATAGA	
		TGA GA A A LITATATCAGCA AGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGG	
CY TS		ATCATGACAAAGTATATCAGCANGGGGGGGGGGGGGGGGG	
	Voe	GATAAAAGICCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGT	849
DD:		GATAAAAGTCCCAACATCCAGTTTTANGGTGTTTTTTTTTT	
· Q7	اق ق	CCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAG	909
ో .Db			
QΥ	.ot	TTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAAC	969
Db^	91	TO A CONTROL OF THE PROPERTY O	972
QY _j	57	0 TACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009	. 17
Db	97		•

RESULT 15 AF113536 LOCUS

```
Homo sapiens MO25 protein mRNA, complete cds.
DEFINITION
ACCESSION
           AF113536
           AF113536.1 GI:6523826
VERSION
KEYWORDS
           Homo sapiens (human)
SOURCE
           Homo sapiens
  ORGANISM
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              (bases 1 to 3466)
REFERENCE
           Jin, W., Shi, J., Ren, S., Gu, J., Fu, S., Huang, Q., Dong, H., Yu, Y.,
  AUTHORS
           Fu, G., Wang, Y., Chen, Z. and Han, Z.
           A novel gene expressed in the human hypothalamus
  TITLE
           Unpublished
  JOURNAL
           2 (bases 1 to 3466)
REFERENCE
           Jin, W., Shi, J., Ren, S., Gu, J., Fu, S., Huang, Q., Dong, H., Yu, Y.,
  AUTHORS
           Fu,G., Wang,Y., Chen,Z. and Han,Z.
           Direct Submission
  TITLE
           Submitted (16-DEC-1998) Chinese National Human Genome Center at
  JOURNAL
           Shanghai, 351 Guo Shoujing Rd., Zhangjiang Hi-Tech Park, Pudong,
           Shanghai 201203, China
                    Location/Qualifiers
FZATURES
                    1. .3466
     source
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /tissue_type="hypothalamus"
                    54. .1079
     CD8
                    /codon start=1
                    /product="MO25 protein"
                    /protein id="AAF14873.1"
                    /db xref="GI:6523827"
                    translation="MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVS"
                    KNLVAMKEILYGTNEKEPQTEAVAQLAQELYNSGLLSTLVADLQLIDFEGKKDVAQIF
                    NNILRPQIGTRTPTVEYICTQQNILFMLLKGYESPEIALNCGIMLRECIRHEPLAKII
                    LWSEQFYDFFRYVEMSTFDIASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYEKLL
                    HSENYVTKRQSLKLLGELLLDRHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFH
                    VFKVFVANPNKTQPILDILLKNQAKLIEFLSKFQNDRTEDEQFNDEKTYLVKQIRDLK
                    ŔPAQQEA"
                                        1070 t
                                689 9
                        606 C
               1101 a
 BASE COUNTY.
 ORIGIN
                                Score 581.6; DB 9; Length 3466;
                         57.4%;
   Query Match
                         74.7%; Pred. No. 7.6e-132;
   Best Local Similarity
                                                    Indels
                                                                 Gaps 1;
                               0; Mismatches 244;
   Matches 747; Conservative
           12 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
 QΥ
              66 TTTGGAAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGCT 125
 Db
           72 ATTTTGGAAAAGCAAGAC------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTCT 129
 Õλ
                                       1111 1111 1111 1111 1111 1111
               125 GTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTCC 185
 dQ
          QУ
                         186 AAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGACA 245
 DΕ
```

QУ	190 GAAGCAGTGGCTCAGCTAGCACAGAACTCTACAGCAGTGGCCTGCTAGCACACTGATA 249
Db	246 GAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGTA 303
ÇУ	250 GCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAAC 309
Db	
Qy	310 ATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCT 369
Db	370 CATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGG 429
Çy Do	370 CATATCCTGTTTATGCTCCTCAAAGGTTTTATGCTCCTCAAAGGTTTTTGTTCATGTTATTGAAAGGGTATGÄATCTCCAGAAATAGCTCTAAATTGTGGA 485
Qy	430 ATTATGCTGAGAGATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAAT 489
Db	486 ATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCCGAAA
QY	490 CAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCC 549
Db	•
QY Ob	550 TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAA 609
SAP	569
ele ele	GSG CAGCATTATGATAGATTTTCAGTGAATATGAGAAGTTACTTCATTCA
ΔĀ.	ACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCC 729
Dp	ACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTAGATAGA
CÄ	730 ATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGG 789.
₽b	790 GATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGT 849
QΥ	
ДЪ	346 GACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAAT 905
СА	850 CCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAG 909
Db	
QY	910 TTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAAC 969
פֿת	970 TACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
ΩĀ	970 TACTTGATTAACAGATCCGAGACTTGAAGATTACGGGGGGGG

.*

71

Search completed: January 6, 2004, 02:34:57 Job time: 3971 secs

GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

January 6, 2004, 00:31:27; Search time 390 Seconds Run on:

(without alignments)

7018.539 Million cell updates/sec

US-10-088-872-1 Title:

1014 Perfect score:

1 atgaaaaaatgcctttgtt.....tgaagaaaacggccccttga 1014 Sequence:

IDENTITY NUC Scoring table:

Gapop 10.0 , Gapext 1.0

2552756 seqs, 1349719017 residues Seamched:

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

N Geneseq 19Jun03:* Datribage 😁

1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:* 2:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:* 3:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:* 4:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*

5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*

6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*

7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*

10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:* 1.1:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:* 12:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:* 13:

/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:* 14:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:* 15:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:* 16:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:* 17:

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*

19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*

20:

/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:* 21:

/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:* 22:

/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:* 23:

/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:* 24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

25:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No Score No Score No Score No Score No Score No Score No No No No No No No N				ર્જ					
1 1014 100.0 1014 22 AAF86462 Human Acute Neuron 2 1014 100.0 1421 22 AAF86462 Human polynucleoti 3 1010.8 99.7 1344 21 AAA27332 Human calcium bind 4 992.8 97.9 2002 22 AAH15879 Human CDNA sequenc 5 770.6 76.0 822 22 AAH65471 Human CDNA sequenc 6 6 684.6 67.5 831 20 AAX39817 Human CDNA CONE (6 684.4 67.5 1191 22 AA160020 Human CDNA CONE (7 684.4 67.5 1191 22 AA160020 Human CDNA CONE (8 592.6 57.5 1026 22 AAC91772 Human ANIC-BP (acu Human DS) 10 582.6 57.5 3849 23 ABV22897 Human prostate exp 11 582.6 57.5 3849 23 ABV22897 Human prostate exp 12 541.6 53.4 1053 22 AAF30688 DNA encoding novel 12 520.2 51.3 833 20 AAX39818 Gastric cancer ass 15 496 48.9 2492 23 AAS89557 DNA encoding novel 14 520.2 51.3 83 20 AAX39818 Gastric cancer ass 15 496 48.9 2492 23 ABV2832 Human gene express 15 496 48.9 2492 23 ABS8031 DNA-encoding novel 14 520.2 51.3 83 20 AAX39818 Gastric cancer ass 15 496 48.9 2492 23 ABS8031 DNA-encoding novel 14 520.2 51.3 83 20 AAX39818 Gastric cancer ass 15 496 48.9 2492 23 ABS8031 DNA-encoding novel 15 496 48.9 2492 23 ABS8031 DNA-encoding novel 16 48.8 24.1 447 24 ABL82285 Human coverian cance 22 244.8 24.1 447 24 ABL82285 Human covarian cance 22 244.8 24.1 447 24 ABL82281 Human covarian cance 24 24.3 24.1 450 24 ABS77084 Frog embryonic gen Human covarian cance 24 210.4 20.7 1474 21 AAC32983 ARA6193 Human covarian cance 25 208.8 20.6 1497 21 AAC42766 AAA314 Arabidopsis thalia 27 200.2 19.7 1032 21 AAC48253 Arabidopsis thalia 27 200.2 19.7 1032 21 AAC48254 ARA6193 Porcine muscular s 20 24 ARA6193 Porcine muscular s 20 25 ARA6193			Q	Query	i ength	DB	TD	Description	
1 1014 100.0 1421 22 AAT58234 Human polynucleoti 3 1010.8 99.7 1344 21 AAA27332 Human cDNA sequenc 4 592.8 97.9 2002 22 AAH15879 Human cDNA sequenc 5 770.6 76.0 822 22 AN65471 Human cDNA clone (6 684.6 67.5 831 20 AAX39817 Gastric cancer ass 6 7 634.4 67.5 1191 22 AA160020 Human polynucleoti 8 582.6 57.5 1026 22 AAC91772 Human polynucleoti 9 582.6 57.5 3849 23 ABV22987 Human prostate exp 10 582.6 57.5 3849 23 ABV22987 Human prostate exp 11 582.6 57.5 3849 23 ABV28822 Human prostate exp 12 541.6 53.4 1053 22 AAF30688 Human acute neuron 13 539.6 53.2 1162 23 AAS89557 DNA encoding novel 13 539.6 53.2 1162 23 AAS89557 DNA encoding novel 14 520.2 51.3 833 20 AAX39818 Gastric cancer ass 15 496 48.9 2492 23 AAS88031 Human gene express 16 367.8 38.2 722 20 AAZ15133 Human gene express 17 362.3 35.8 2231 23 ABL07151 Drosophila melanog 18 362.8 35.8 4231 23 ABL07150 Drosophila melanog 19 283.8 28.5 690 24 ABS77084 Frog embryonic gen 20 246.4 42.3 435 24 ABL82285 Human ovarian canc 21 244.8 24.1 447 24 ABL82285 Human ovarian canc 22 244.3 24.1 450 24 ABL81975 Human ovarian canc 23 120.3 20.8 762 24 ABS76784 Frog embryonic gen 24 310.4 20.7 1474 21 AAC23983 Arabidopsis thalia 25 208.8 20.6 1497 21 AAC40181 Arabidopsis thalia 26 200.2 19.7 918 21 AAC40181 Arabidopsis thalia 27 380.2 19.7 1032 21 AAC42766 Arabidopsis thalia 28 195 19.2 387 24 ABN3983 Gene #481 used to 29 195 19.2 387 24 ABN9824 Human ovarian canc 30 169.9 16.7 722 24 AAS61999 Porcine muscular s 31 166.6 16.4 700 24 AAS61999 Porcine muscular s 32 163.8 16.2 481 25 ABZ19574 Human gene express 34 161.2 15.9 1515 21 AAC42766 Arabidopsis thalia 35 2163.8 16.2 481 25 ABZ19574 Human gene express 36 161.2 15.9 1515 21 AAC40766 Arabidopsis thalia 36 163.4 16.1 300 20 AAS88930 DNA encoding novel 36 163.4 16.1 300 20 AAS88930 DNA encoding novel 37 147.2 14.5 464 21 AAC50415 AAS61993 Porcine muscular s 38 161.2 15.9 1515 21 AAC40766 Arabidopsis thalia 39 107.6 10.6 1149 23 AAS88930 DNA encoding novel 40 107.6 10.6 3279 23 AAS8930 DNA encoding novel 41 65.6 6.4 487 22 AA	.4	No.	Score	Maccii	nengen				
2 10104 100.0 1421 22 AA158234 Human polynucleoting 3 1010.8 99.7 1344 21 AAA27332 Human calcium bind 4 592.8 97.9 2002 22 AAH15879 Human cDNA sequence 5 770.6 76.0 822 22 AAH15879 Human cDNA clone (6 684.6 67.5 813 20 AAX39817 Gastric cancer ass 8 582.6 57.5 1026 22 AC91772 Human ANIC-BP (acu 9 592.6 57.5 3281 24 ABK13127 Human secretory polynomial 582.6 57.5 3849 23 ABV22987 Human prostate exp 11 582.6 57.5 3849 23 ABV2822 Human prostate exp 12 541.6 53.4 1053 22 AAF30688 Human prostate exp 12 541.6 53.4 1053 22 AAF30688 Human prostate exp 12 541.6 53.4 1053 22 AAF30688 Human prostate exp 13 539.6 53.2 1162 23 AAS89557 DNA encoding novel 6 14 520.2 51.3 833 20 AAX39818 Gastric cancer ass 15 496 48.9 2492 23 AAS88031 DNA encoding novel 16 387.8 38.2 722 20 AAZ15133 Human gene express 16 387.8 38.2 722 20 AAZ15133 Human gene express 17 362.3 35.8 2231 23 ABL07151 Drosophila melanog 19 283.8 28.5 690 24 ABS77084 Human ovarian canc 10 246.4 24.3 435 24 ABL82285 Human ovarian canc 22 344.2 24.1 447 24 ABL82285 Human ovarian canc 23 420.3 20.8 762 24 ABL82291 Human ovarian canc 24 310.4 20.7 1474 21 AAC32983 Arabidopsis thalia 25 208.9 20.6 1497 21 AAC42766 Arabidopsis thalia 26 200.2 19.7 1938 21 AAC42766 Arabidopsis thalia 3 166.6 16.4 700 24 ABS76784 Frog embryonic gen 29 195 19.2 387 24 ABS9993 Gene #481 used to 1.6 2.3 1.6 2.4 1.5 46 2.4 1.5 46 2.4 1.5 46 2.4 1.5 46 2.4 1.5 46 2.4 1.5 46 2.4 1.5 46 2.4 1.5 46 2.4 1.5 46 2.5 4.5 4.5 4.5 4.5 4.5 4.5 4.5 4.5 4.5 4		1	1014	100.0	1014	22	AAF86462		
3 1010.8 99.7 1344 21 AAA27332 Human Collam blnd 4 592.8 97.9 2002 22 AAH15879 Human collam sequence					1421	22	AAI58234		
## 592.8 97.9 2002 22 AAH15879						21	AAA27332		
5 770.6 76.0 822 22 AAR05471 Human clark Clone 6 684.6 67.5 831 20 AAX39817 Gastric cancer ass 8 582.6 57.5 1091 22 AAC901702 Human polynucleoti 9 582.6 57.5 3281 24 ABC91772 Human secretory po 10 582.6 57.5 3849 23 ABV28822 Human prostate exp 11 582.6 57.5 3849 23 ABV28822 Human prostate exp 12 541.6 53.4 1053 22 AAF30688 Human acute neuron 13 539.6 53.2 1162 23 AAS99557 Gastric cancer ass 15 496 48.9 2492 23 AAS89051 DNA encoding novel 16 387.8 38.2 722 20 AAX39818 DNA encoding novel 16 387.8 38.2 722 20 AAX19813 DNA encoding novel 18 362.3 35.8 2231 23 ABL07150 Drosophila melanog 19 283.8 28.5 690 24 ABS77084 Fuman ovarian cance 21 444.8 24.1 447 24 ABL82921 Human ovarian cance 22 544.2 24.1 450 24 ABL82921 Human ovarian cance 23 520.8 762 24 ABS77084 Frog embryonic gen 24 210.4 20.7 1474 21 AAC32983 Arabidopsis thalia 25 208.9 20.6 1497 21 AAC42766 Arabidopsis thalia 26 200.2 19.7 1032 21 AAC48253 Gene #481 used to 29 195 19.2 387 24 ABL93983 Gene #481 used to 29 195 19.2 387 24 ABS61992 Porcine muscular s 20 163.4 16.1 300 20 AAX1652 Human ovarian cance 21 163.8 16.7 722 24 AAS61992 Porcine muscular s 21 163.8 16.2 481 25 ABZ19574 Group III cDNA can 32 163.4 16.1 300 20 AAX1652 Human ovarian cance 22 163.8 16.7 722 24 AAS61992 Porcine muscular s 21 163.8 16.2 481 25 ABZ19574 Group III cDNA can 32 163.4 16.1 300 20 AAX1652 Human ovarian cance 24 163.8 16.2 481 25 ABZ19574 Group III cDNA can 32 163.4 16.1 300 20 AAX14552 Human ovarian cance 24 163.8 16.2 481 25 ABZ19574 Group III cDNA can 32 163.4 16.1 300 20 AAX14552 Human ovarian cance 24 163.8 16.2 481 25 ABZ19574 Group III cDNA can 32 163.4 16.1 300 20 AAX14552 Human ovarian cance 24 163.8 16.2 481 25 ABZ19574 Group III cDNA can 32 163.4 16.1 300 20 AAX14552 Human ovarian cance 24 163.8 16.2 481 25 ABZ19574 Group III cDNA can 32 163.4 16.1 300 20 AAX14552 Human ovarian cance 24 163.8 16.2 481 25 ABZ19574 Group III cDNA can 32 163.4 16.1 300 20 AAX14552 Human ovarian cance 24 163.8 16.2 481 25 ABZ19574 Group III cDNA can 32 163.4 16.1 300 20 AAX14552 Huma						22	AAH15879		
C 7 684.6 67.5 831 20 AAX39817 Gastric cancer associated by the cancer					822	22	AAH05471		
C 7 634.4 67.5 1191 22 AA160020 Human polyhulcelocit Human ANIC-BP (acu Human ANIC-BP) (acu Human Prostate exp Human prostate exp Human prostate exp Human prostate exp Human acute neuron AAS9957 DNA encoding novel AAS9957 DNA encoding novel AAS9957 DNA encoding novel Gastric cancer ass DNA encoding novel Human acute neuron AAX39818 Gastric cancer ass DNA encoding novel Human gene express DNA encoding novel Human gene express DNA encoding novel Human gene express DNA encoding novel Human Gastric cancer ass DNA encoding novel AAX39818 Gastric cancer ass DNA encoding novel Human gene express DNA encoding novel Human Gastric Cancer ass DNA encoding novel ABB61975 Human Ovarian cance Human Ovarian cance Human Ovarian cance Human Ovarian cance ABB61975 Human Ovarian cance Human Ovarian cance ABB61975 Human Ovarian cance ABB61975 Human Ovarian cance Human Ovarian Cancer Canc					831	20	AAX39817		
8 582.6 57.5 1026 22 AAC91772 Human secretory po 582.6 57.5 3281 24 ABK13127 Human secretory po 10 582.6 57.5 3281 24 ABK13127 Human prostate exp 11 582.6 57.5 3849 23 ABV22987 Human prostate exp 12 541.6 53.4 1053 22 AAF30688 Human prostate exp 13 59.6 53.2 1162 23 AAS39688 Human acute neuron DNA encoding novel Gastric cancer ass DNA encoding novel 16 367.8 38.2 722 20 AAZ15133 DNA encoding novel 16 367.8 38.2 722 20 AAZ15133 Human gene express 17 362.3 35.8 2231 23 ABL07150 Drosophila melanog 19 283.8 28.5 690 24 ABS77084 Frog embryonic gen 19 283.8 28.5 690 24 ABS77084 Frog embryonic gen 19 246.4 24.3 435 24 ABL82285 Human ovarian cance 20 246.4 24.1 447 24 ABL82285 Human ovarian cance 21 244.8 24.1 447 24 ABL82285 Human ovarian cance 22 244.2 24.1 447 24 ABL82921 Human ovarian cance 23 244.0 24.1 447 24 ABL82921 Human ovarian cance 24 24.1 20.7 1474 21 AAC30181 Arabidopsis thalia 25 208.8 20.6 1497 21 AAC40181 Arabidopsis thalia 25 208.8 20.6 1497 21 AAC40766 Arabidopsis thalia 25 208.8 20.6 1497 21 AAC42766 Arabidopsis thalia 26 28 195 19.2 387 24 ABN93983 Gene #481 used to 1.07.6 16.6 16.4 700 24 AAS61992 Porcine muscular s 21 63.8 16.2 481 25 AB219574 Group III cDNA can 4.07.6 16.6 16.4 700 24 AAS61993 Porcine muscular s 2.06 153.4 15.1 737 23 AAS79449 Porcine muscular s 2.06 153.4 15.1 737 23 AAS79449 Porcine muscular s 2.06 153.4 15.1 737 23 AAS79449 Porcine muscular s 2.06 153.4 15.1 737 23 AAS79449 Porcine muscular s 2.06 153.4 15.1 737 23 AAS79449 Porcine muscular s 2.06 153.4 15.1 737 23 AAS79449 Porcine muscular s 2.06 153.4 15.1 737 23 AAS79449 Porcine muscular s 2.06 153.4 15.1 737 23 AAS89830 DNA encoding novel 40 107.6 10.6 12.79 23 AAS89559 DNA encoding novel 41 65.6 6.5 432 24 AB107116 Human cDNA clone (10.07 10.07	C				1191	22	AA160020		
9 582.6 57.5 3281 24 ABK13127 Human prostate exp 10 582.6 57.5 3849 23 ABV22987 Human prostate exp 11 582.6 57.5 3849 23 ABV2287 Human prostate exp 12 541.6 53.4 1053 22 AAF30688 Human acute neuron 13 539.6 53.2 1162 23 AAS89557 DNA encoding novel 14 520.2 51.3 833 20 AAX39818 Gastric cancer ass 15 496 48.9 2492 23 AAS88031 DNA encoding novel 16 367.8 38.2 722 20 AAZ15133 Human gene express 17 367.3 35.8 2231 23 ABL07151 Drosophila melanog 18 362.8 35.8 4231 23 ABL07151 Drosophila melanog 19 283.8 28.5 690 24 ABS77084 Frog embryonic gen 19 284.8 24.1 447 24 ABL82921 Human ovarian canc 21 044.8 24.1 447 24 ABL82921 Human ovarian canc 22 046.4 24.3 435 24 ABL81975 Human ovarian canc 23 120.8 24.1 450 24 ABL81975 Human ovarian canc 24 210.4 20.7 1474 21 AAC42983 Arabidopsis thalia 25 208.9 20.6 1497 21 AAC40181 Arabidopsis thalia 26 200.2 19.7 918 21 AAC42766 Arabidopsis thalia 26 200.2 19.7 1032 21 AAC48253 Arabidopsis thalia 27 300.2 19.7 1032 21 AAC48253 Arabidopsis thalia 28 195 19.2 387 24 ABN93983 Gene #481 used to 29 195 19.2 387 24 ABS61992 Porcine muscular s 20 163.8 16.2 481 25 ABZ19574 Human gene express 34 161.2 15.9 1515 21 AAC5915 Arabidopsis thalia 35 163.4 16.1 300 20 AAZ1552 Human gene express 36 163.4 16.1 300 20 AAZ1552 Human gene express 37 1A7.2 14.5 464 21 AAN93824 ARAB0393 Porcine muscular s 38 133.2 13.1 615 22 AAN98824 Arabidopsis thalia 39 163.4 16.1 300 20 AAZ1552 Human gene express 30 169.9 16.7 722 24 AAS61993 Porcine muscular s 30 169.9 16.7 722 24 AAS61993 Porcine muscular s 31 163.6 15.4 861 24 ABN98824 Arabidopsis thalia 32 163.8 16.2 481 25 ABZ19574 Human gene express 34 161.2 15.9 1515 21 AAC46721 Human gene express 35 163.4 16.1 300 20 AAZ1552 Human gene express 36 133.2 13.1 615 22 AAH07116 Human coMA clone (36 153.4 16.1 300 20 AAZ16592 Porcine muscular s 37 1A7.2 14.5 464 21 AAC46721 Porcine muscular s 38 133.2 13.1 615 22 AAH07116 Human coMA clone (38 133.2 13.1 615 22 AAH07116 Human coMA clone (39 107.6 10.6 1279 23 AAS89559 DNA encoding novel Human DNA encoding novel Hu	•				1026	22	AAC91772		
10 582.6 57.5 3849 23 ABV22987 Human prostate exp 11 582.6 57.5 3849 23 ABV28822 Human prostate exp 12 541.6 53.4 1053 22 AAF30688 Human acute neuron 13 539.6 53.2 1162 23 AAF30688 Human acute neuron 14 520.2 51.3 833 20 AAX39818 Gastric cancer ass 15 496 48.9 2492 23 AAF38031 DNA encoding novel 16 387.8 38.2 722 20 AAZ15133 Human gene express 17 362.3 35.8 2231 23 ABL07151 Drosophila melanog 18 362.8 35.8 4231 23 ABL07151 Drosophila melanog 19 263.8 28.5 690 24 ABF37084 Frog embryonic gen 20 245.4 24.3 435 24 ABL82285 Human ovarian canc 21 444.8 24.1 447 24 ABL82921 Human ovarian canc 22 344.2 24.1 450 24 ABL81975 Human ovarian canc 23 220.3 20.8 762 24 ABF37684 Frog embryonic gen 24 310.4 20.7 1474 21 AAC32983 Arabidopsis thalia 25 208.9 20.6 1497 21 AAC40181 Arabidopsis thalia 26 200.2 19.7 918 21 AAC42766 Arabidopsis thalia 27 300.2 19.7 1032 21 AAC48253 Gene 481 used to 29 195 19.2 387 24 ABN93983 Gene 481 used to 29 195 19.2 387 24 ABS61992 Porcine muscular s 30 169.9 16.7 722 24 AAS61992 Porcine muscular s 31 166.6 16.4 700 24 AAS61992 Porcine muscular s 32 163.8 16.2 481 25 ABS219574 Group III cDNA can 33 163.4 16.1 300 20 AAZ14552 ARB09834 ARAbidopsis thalia 34 161.2 15.9 1515 21 AAC4571 ARB093 Porcine muscular s 36 163.4 16.1 300 20 AAZ14552 ARB09393 Human 36 163.4 16.1 300 20 AAZ14552 ARB09393 Human 37 153.4 15.1 737 23 AAS7949 Porcine muscular s 38 133.2 13.1 615 22 AAH07116 Human 39 107.6 10.6 1149 23 AAS88030 DNA encoding novel 40 07.6 10.6 3279 23 AAS89559 Human 40 107.6 10.6 1149 23 AAS88030 DNA encoding novel 41 65.6 6.4 487 22 AAN67107 Human 42 65 6.4 487 22 AAN67107 Human 43 65 6.4 487 22 AAN67107 Human 44 65 6.4 487 22 AAN67107 Human 45 AAS213110 Human 46 Human 47 Human 48 DNA encoding novel 48 Human 49 Human 40 Human 41 Human 40 Human 41 Human 41 Human 41 Human 41 Human 4					3281	24	ABK13127		
11 582.6 57.5 3849 23 ABV28822 Human prostate exp 12 541.6 53.4 1053 22 AAF30688 Human acute neuron 13 539.6 53.2 1162 23 AAS89557 DNA encoding novel 14 520.2 51.3 833 20 AAS88031 DNA encoding novel 15 496 48.9 2492 23 AAS88031 DNA encoding novel 16 387.8 38.2 722 20 AAZ15133 Human gene express 17 362.3 35.8 2231 23 ABL07151 Drosophila melanog 18 362.8 35.8 4231 23 ABL07150 Drosophila melanog 19 283.8 28.5 690 24 ABS77084 Frog embryonic gen 20 245.4 24.3 435 24 ABL82285 Human ovarian canc 21 344.8 24.1 447 24 ABL82295 Human ovarian canc 22 344.2 24.1 450 24 ABL81975 Human ovarian canc 23 320.3 20.8 762 24 ABS76784 Frog embryonic gen 24 210.4 20.7 1474 21 AAC32983 Arabidopsis thalia 25 208.9 20.6 1497 21 AAC40181 Arabidopsis thalia 26 200.2 19.7 918 21 AAC42766 Arabidopsis thalia 27 300.2 19.7 1032 21 AAC48253 Gene #481 used to 28 195 19.2 387 24 ABL8939 Gene #481 used to 29 195 19.2 387 24 ABL66143 Lung cancer relate 30 169.9 16.7 722 24 AAS61992 Porcine muscular s 31 166.6 16.4 700 24 AAS61993 Porcine muscular s 32 163.8 16.2 481 25 ABZ19574 Group III cDNA can 33 163.4 16.1 300 20 AAZ14552 ABZ19574 Group III cDNA can 34 161.2 15.9 1515 21 AAC46721 ARABIDOPS thalia 35 153.4 161.1 737 23 AAS79449 Porcine muscular s 36 153.4 15.1 737 23 AAS9949 Porcine muscular s 36 153.4 15.1 737 23 AAS9949 Porcine muscular s 37 147.2 14.5 464 21 AAC46721 Ea mays DNA fragm 38 133.2 13.1 615 22 AAH07116 Human cDNA clone (57.5	3849	23	ABV22987		
12 541.6 53.4 1053 22 AAF30688 Human acute neuron DNA encoding novel 639.6 53.2 1162 23 AAS89557 DNA encoding novel 14 520.2 51.3 833 20 AAX39818 Gastric cancer ass 15 496 48.9 2492 23 AAS88031 DNA encoding novel 16 387.8 38.2 722 20 AAZ15133 Human gene express 2.7 362.3 35.8 2231 23 ABL07151 Drosophila melanog 18 362.8 35.8 4231 23 ABL07150 Drosophila melanog 19 283.8 28.5 690 24 ABS77084 Frog embryonic gen 20 246.4 24.3 435 24 ABL82285 Human ovarian canc 21 244.8 24.1 447 24 ABL82285 Human ovarian canc 22 244.1 450 24 ABL81975 Human ovarian canc 22 24.2 24.1 450 24 ABL81975 Human ovarian canc 24 210.4 20.7 1474 21 AAC32983 Arabidopsis thalia 25 208.8 20.6 1497 21 AAC40181 Arabidopsis thalia 26 200.2 19.7 918 21 AAC42766 Arabidopsis thalia 27 M300.2 19.7 1032 21 AAC48253 Arabidopsis thalia 27 M300.2 19.7 1032 21 AAC48253 Arabidopsis thalia 30 169.8 16.7 722 24 AAS61993 Gene #481 used to 29 195 19.2 387 24 ABS93983 Gene #481 used to 29 195 19.2 387 24 ABS93983 Gene #481 used to 29 163.8 16.7 722 24 AAS61993 Porcine muscular s 20 163.8 16.2 481 25 ABZ19574 Group III cDNA can 32 163.4 16.1 300 20 AAZ14552 Human gene express 34 161.2 15.9 1515 21 AAC50415 Arabidopsis thalia 35 156 15.4 861 24 ABN93824 Arabidopsis thalia 36 153.4 15.1 737 23 AAS79449 DNA encoding novel 26 27 14.5 464 21 AAC40721 ARB09305 DNA fragm 38 133.2 13.1 615.2 AAM97116 Human CDNA clone (DNA encoding novel 10.6 10.6 3279 23 AAS88030 DNA encoding novel 10.6 10.6 3279 23 AAS88030 DNA encoding novel 10.6 6.5 6.4 487 22 AAI98879 Human GDP-mannose 41 65.6 6.4 487 22 AAI98879 Human DAAC4070 Human CDNA clone (DNA encoding novel 24 65 6.4 487 22 AAI98879 Human Excreteory re 244 53.6 5.3 254 25 ABX31310 Human GDP-mannose 41 AAB0424 53.6 5.3 254 25 ABX31310 Human GDP-mannose 41 AAB0424 53.6 5.3 254 25 ABX31310 Human GDP-mannose 41 AAB0424 53.6 5.3 254 25 ABX31310 Human GDP-mannose 41 AAB0424 53.6 5.3 254 25 ABX31310 Human GDP-mannose 41 AAB0424 53.6 5.3 254 25 ABX31310 Human GDP-mannose 41 AAB0424 53.6 5.3 254 25 ABX31310 Human GDP-mannose 41 AAB0424 5					3849	23	ABV28822	· · · · · · · · · · · · · · · · · · ·	
C 14 520.2 51.3 833 20 AAX39818 Gastric cancer ass DNA encoding novel 48.9 2492 23 AAS88031 DNA encoding novel 38.7 496 48.9 2492 23 AAS88031 DNA encoding novel 48.9 2492 23 AAS88031 Human gene express 27 362.3 35.8 2231 23 ABL07151 Drosophila melanog 28.5 690 24 ABS77084 Frog embryonic gen 20 246.4 24.3 435 24 ABL82285 Human ovarian canc 21 444.8 24.1 447 24 ABL82285 Human ovarian canc 22 444.8 24.1 447 24 ABL82921 Human ovarian canc 24 42 42.3 42.1 450 24 ABL81975 Human ovarian canc 24 42.2 44.1 450 24 ABL81975 Human ovarian canc 24 42.0 24.1 450 24 ABL81975 Human ovarian canc 24 42.0 20.8 762 24 ABS76784 Frog embryonic gen 24 21.0 4 20.7 1474 21 AAC32983 Arabidopsis thalia 25 208.8 20.6 1497 21 AAC40181 Arabidopsis thalia 26 200.2 19.7 918 21 AAC40181 Arabidopsis thalia 27 800.2 19.7 1032 21 AAC48253 Arabidopsis thalia 27 800.2 19.7 1032 21 AAC48253 Arabidopsis thalia 30 169.3 16.7 722 24 ABS1993 Gene 4481 used to 29 195 19.2 387 24 ABL66143 Lung cancer relate 30 169.8 16.7 722 24 AAS61992 Porcine muscular s 31 166.6 16.4 700 24 AAS61992 Porcine muscular s 32 163.8 16.2 481 25 ABZ19574 Human gene express 34 161.2 15.9 1515 21 AAC50415 Arabidopsis thalia 35 15.4 861 24 ABN98824 DNA encoding novel 26 153.4 15.1 737 23 AAS79449 DNA encoding novel 27 147.2 14.5 464 21 AAC46721 Zea mays DNA fragm 38 133.2 13.1 615 22 AAH07116 Human CDNA clone (DNA encoding novel 41 65.6 6.5 432 24 ABN98895 DNA encoding novel 41 65.6 6.4 487 22 AAN98107 Human excretory re 44 65 6.4 487 22 AAN98107 Human excretory re 44 65 6.4 487 22 AAN98107 Human excretory re 44 65 6 6.4 487 22 AAN98107 Human excretory re 44 65 6 6.4 487 22 AAN98107 Human excretory re 44 65 6 6.4 487 22 AAN98107 Human excretory re 44 65 6 6.4 487 22 AAN98107 Human excretory re 44 65 6 6.4 487 22 AAN98107 Human excretory re 44 65 6 6.4 487 22 AAN98107 Human excretory re 44 65 6 6.4 487 22 AAN98107 Human excretory re 44 65 6 6.4 487 22 AAN98107 Human excretory re 44 65 6 6.4 487 22 AAN981310 Human BDP-mannose					1053	22	AAF30688		
C 14 520.2 51.3 833 20 AAX39818 Gastric Cancel ass 15 496 48.9 2492 23 AAS88031 DNA encoding novel 16 387.8 38.2 722 20 AAX15133 Human gene express 17 362.3 35.8 2231 23 ABL07151 Drosophila melanog 18 362.8 35.8 4231 23 ABL07150 Drosophila melanog 19 283.8 28.5 690 24 ABS77084 Frog embryonic gen 20 246.4 24.3 435 24 ABL82285 Human ovarian canc 21 344.8 24.1 447 24 ABL82285 Human ovarian canc 22 344.0 24.1 450 24 ABL81975 Human ovarian canc 23 342.0 20.8 762 24 ABS76784 Frog embryonic gen 24 210.4 20.7 1474 21 AAC32983 Arabidopsis thalia 25 208.8 20.6 1497 21 AAC40181 Arabidopsis thalia 26 200.2 19.7 918 21 AAC42766 Arabidopsis thalia 27 N00.2 19.7 1032 21 AAC48253 Arabidopsis thalia 27 N00.2 19.7 1032 21 AAC48253 Arabidopsis thalia 28 195 19.2 387 24 ABL66143 Gene #481 used to 29 195 19.2 387 24 ABS61992 Porcine muscular s 31 166.6 16.4 700 24 AAS61993 Porcine muscular s 32 163.8 16.2 481 25 ABZ19574 Group III cDNA can 33 163.4 16.1 300 20 AAZ14552 Arabidopsis thalia 35 163.4 16.1 300 20 AAZ14552 Arabidopsis thalia 35 15.6 15.4 861 24 ABN98824 Arabidopsis thalia 36 153.4 15.1 737 23 AAS8030 DNA encoding novel 37 147.2 14.5 464 21 AAC46721 Zea mays DNA fragm 133.2 13.1 615 22 AAH07116 Human cDNA clone (DNA encoding novel 40 107.6 10.6 3279 23 AAS8030 DNA encoding novel 41 65.6 6.5 432 24 ABN98879 Human cRF3054 cDNA 42 65 6.4 487 22 AAI64066 Human GDP-mannose 44 53.6 5.3 254 25 ABX31310 Human GDP-mannose 44 53.6 5.3 254 25 ABX31310 Human GDP-mannose					1162	23	AAS89557		
15 496 48.9 2492 23 AAS88031 16 387.8 38.2 722 20 AAZ15133 17 362.3 35.8 2231 23 ABL07150 18 362.3 35.8 4231 23 ABL07150 19 283.8 28.5 690 24 ABS77084 20 246.4 24.3 435 24 ABL82285 21 444.8 24.1 447 24 ABL82285 22 32.0 32 20.8 762 24 ABS81975 23 320.3 20.8 762 24 ABS81975 24 210.4 20.7 1474 21 AAC32983 25 208.8 20.6 1497 21 AAC40181 26 200.2 19.7 918 21 AAC42766 27 300.2 19.7 918 21 AAC42766 28 195 19.2 387 24 ABS93983 29 20 24 ABS61992 30 169.9 16.7 722 24 AAS61992 31 166.6 16.4 700 24 AAS61992 32 163.8 16.2 481 25 ABZ19574 33 163.4 16.1 300 20 AAZ14552 34 161.2 15.9 1515 21 AAC46721 35 156 15.4 861 24 ABN98824 36 153.4 15.1 737 23 AAS79449 37 147.2 14.5 464 21 AAC46721 38 133.2 13.1 615 22 AAH07116 39 107.6 10.6 3279 23 AAS88030 30 Aego ABS7804 ABN9805 40 107.6 10.6 3279 23 AAS89559 41 65.6 6.5 432 24 ABN78107 42 ABS78107 42 ABS98559 44 BDN98824 45 ABR98879 46 BN98879 47 BN98824 48 BN98824 48 ARAB1dopsis thalia 49 ARAB1993 40 AR	а				833	20	AAX39818		
16 387.8 38.2 722 20 AAZ15133 Human gene express 17 362.3 35.8 2231 23 ABL07151 Drosophila melanog 18 362.8 35.8 4231 23 ABL07150 Drosophila melanog 19 283.8 28.5 690 24 ABS77084 Frog embryonic gen 20 246.4 24.3 435 24 ABL82285 Human ovarian canc 21 644.8 24.1 447 24 ABL82921 Human ovarian canc 22 244.8 24.1 450 24 ABL81975 Human ovarian canc 24 210.4 20.7 1474 21 AAC32983 Arabidopsis thalia 25 208.8 20.6 1497 21 AAC40181 Arabidopsis thalia 26 200.2 19.7 918 21 AAC42766 Arabidopsis thalia 27 300.2 19.7 1032 21 AAC48253 Arabidopsis thalia 28 195 19.2 387 24 ABS93983 Gene #481 used to 29 195 19.2 387 24 ABS6193 Gene #481 used to 29 195 19.2 387 24 AAS61992 Porcine muscular s 32 163.8 16.2 481 25 ABZ19574 Group III cDNA can 32 163.4 16.1 300 20 AAZ14552 Human gene express 34 161.2 15.9 1515 21 AAC50415 Arabidopsis thalia 35 156 15.4 861 24 ABN98824 Arabidopsis thalia 35 156 15.4 861 24 ABN98824 Arabidopsis thalia 36 153.4 15.1 737 23 AAS79449 DNA encoding novel 37 147.2 14.5 464 21 AAC46721 Zea mays DNA fragm 40 107.6 10.6 3279 23 AAS88030 DNA encoding novel 40 107.6 10.6 3279 23 AAS88030 DNA encoding novel 41 65.6 6.4 487 22 AAI98879 Human CRF3054 cDNA Human GDP-mannose 44 53.6 5.3 254 25 ABX31310 Human GDP-mannose	_			48.9	2492	23	AAS88031		
27 362.3 35.8 2231 23 ABL07151 Drosophila melanog 18 362.8 35.8 4231 23 ABL07150 Drosophila melanog 19 283.8 28.5 690 24 ABS77084 Frog embryonic gen 20 245.4 24.3 435 24 ABL82285 Human ovarian canc 21 444.8 24.1 447 24 ABL82921 Human ovarian canc 22 444.2 24.1 450 24 ABL81975 Human ovarian canc 23 210.0 3 20.8 762 24 ABS76784 Frog embryonic gen 24 210.4 20.7 1474 21 AAC32983 Arabidopsis thalia 25 208.8 20.6 1497 21 AAC40181 Arabidopsis thalia 26 200.2 19.7 918 21 AAC42766 Arabidopsis thalia 27 200.2 19.7 1032 21 AAC48253 Arabidopsis thalia 28 29 195 19.2 387 24 ABN93983 Gene #481 used to 29 195 19.2 387 24 ABS61992 Porcine muscular s 30 169.8 16.7 722 24 AAS61992 Porcine muscular s 31 166.6 16.4 700 24 AAS61992 Porcine muscular s 32 163.8 16.2 481 25 ABZ19574 Group III cDNA can 33 163.4 16.1 300 20 AAZ14552 Human gene express 34 161.2 15.9 1515 21 AAC50415 Arabidopsis thalia 35 156 15.4 861 24 ABN98824 Arabidopsis thalia 36 153.4 15.1 737 23 AAS79449 DNA encoding novel 37 147.2 14.5 464 21 AAC46721 Zea mays DNA fragm 38 133.2 13.1 615 22 AAH07116 Human cDNA clone (39 107.6 10.6 1149 23 AAS88030 DNA encoding novel 40 107.6 10.6 3279 23 AAS89559 Human excretory re 41 65.6 6.4 487 22 AAI98879 Human excretory re 42 65 6.4 487 22 AAI98879 Human cDP-mannose 43 65 6.4 487 22 AAI98879 Human bladder rela 44 53.6 53.2 254 25 ABX31310 Human GDP-mannose						20	AAZ15133		
18				35.8	2231	23	ABL07151		
19 283.8 28.5 690 24 ABS77084 Frog embryonic general cances and the second cances and the second cances and ca						23	ABL07150		
20 246.4 24.3 435 24 ABL82285 Human ovarian cance 21 244.8 24.1 450 24 ABL82921 Human ovarian cance 22 244.2 24.1 450 24 ABL81975 Human ovarian cance 23 22.0.3 20.8 762 24 ABS76784 Frog embryonic gen 24 210.4 20.7 1474 21 AAC32983 Arabidopsis thalia 25 208.8 20.6 1497 21 AAC40181 Arabidopsis thalia 26 200.2 19.7 918 21 AAC42766 Arabidopsis thalia 27 200.2 19.7 1032 21 AAC42766 Arabidopsis thalia 28 195 19.2 387 24 ABN93983 Gene #481 used to 29 195 19.2 387 24 ABS61992 Porcine muscular s 20 169.8 16.7 722 24 AAS61993 Porcine muscular s 21 166.6 16.4 700 24 AAS61993 Porcine muscular s 22 163.8 16.2 481 25 ABZ19574 Group III cDNA can 23 163.4 16.1 300 20 AAZ14552 Human gene express 24 161.2 15.9 1515 21 AAC50415 Arabidopsis thalia 25 156 15.4 861 24 ABN98824 Arabidopsis thalia 26 26 153.4 15.1 737 23 AAS79449 DNA encoding novel 27 147.2 14.5 464 21 AAC46721 Eea mays DNA fragm 28 133.2 13.1 615 22 AAH07116 Human cDNA clone (29 107.6 10.6 1249 23 AAS88030 DNA encoding novel 29 107.6 10.6 3279 23 AAS88030 DNA encoding novel 20 107.6 10.6 3279 23 AAS88030 DNA encoding novel 21 107.6 10.6 3279 23 AAS88030 DNA encoding novel 22 143.6 487 22 AAI98879 Human excretory re 23 144.8 53.6 5.3 254 25 ABX31310 Human Bladder rela 24 144.8 25 ABX31310 Human Bladder rela 25 150.6 5.3 254 25 ABX31310 Human secreted pro				28.5	690	24	ABS77084		
21				24.3	435	24	ABL82285		
22 0.44 .3 24.1 450 24 ABL81975 Human OvaFlan Cane 23 0.20.8 762 24 ABS76784 Frog embryonic gen 24 210.4 20.7 1474 21 AAC32983 Arabidopsis thalia 25 208.8 20.6 1497 21 AAC40181 Arabidopsis thalia 26 200.2 19.7 918 21 AAC42766 Arabidopsis thalia 27 300.2 19.7 1032 21 AAC48253 Arabidopsis thalia 6 28 195 19.2 387 24 ABN93983 Gene #481 used to C 29 195 19.2 387 24 ABL66143 Lung cancer relate 30 169.8 16.7 722 24 AAS61992 Porcine muscular s 31 166.6 16.4 700 24 AAS61993 Porcine muscular s 32 163.8 16.2 481 25 ABZ19574 Grou				24.1	447	24	ABL82921		
23				24.1	450	24	ABL81975		
24				20.8	762	24	ABS76784		
25 208.8 20.6 1497 21 AAC40181 Arabidopsis thalia Lung cancer relate Porcine muscular s Arabidopsis thalia				20.7	1474	21	AAC32983		
26 200.2 19.7 918 21 AAC42766 27 200.2 19.7 1032 21 AAC48253				20.6	1497	21	AAC40181		
27 300.2 19.7 1032 21 AAC48253 C 28 195 19.2 387 24 ABN93983 C 29 195 19.2 387 24 ABL66143 Dung cancer relate 30 169.8 16.7 722 24 AAS61992 Porcine muscular s 31 166.6 16.4 700 24 AAS61993 B 163.8 16.2 481 25 ABZ19574 B 161.2 15.9 1515 21 AAC50415 B 161.2 15.9 1515 21 AAC50415 B 161.2 15.9 1515 21 AAC50415 B 163.4 16.1 737 23 AAS79449 B 163.4 15.1 737 23 AAS79449 B 163.4 15.1 737 23 AAS79449 B 163.2 13.1 615 22 AAH07116 B 163.2 13.1 615 22 AAH07116 B 163.3 133.2 13.1 615 22 AAH07116 B 165.6 6.5 432 24 ABN78107 B 165.6 6.5 432 24 ABN78107 B 165.6 6.4 487 22 AAI98879 B 167.6 10.6 1149 23 AAS8930 B 167.6 10.6 1149 23 AAS89559 B 167.6 168 169 169 169 169 169 169 169 169 169 169				19.7	918	21	AAC42766		
C 28 195 19.2 387 24 ABN93983 C 29 195 19.2 387 24 ABL66143 D 169.8 16.7 722 24 AAS61992 D 166.6 16.4 700 24 AAS61993 D 163.8 16.2 481 25 ABZ19574 D 163.4 16.1 300 20 AAZ14552 D 164.2 15.9 1515 21 AAC50415 D 156 15.4 861 24 ABN98824 D 153.4 15.1 737 23 AAS79449 D 163.4 15.1 737 23 AAS79449 D 163.2 13.1 615 22 AAH07116 D 163.2 13.1 615 22 AAH07116 D 163.3 107.6 10.6 1149 23 AAS8030 D 107.6 10.6 3279 23 AAS89559 D 107.6 10.6 3279 24 ABN78107 D 107.6 10.6 3279 25 AAI64066 D 107.6 10.6 3279 24 AAI64066 D 107.6 10.6 3279 25 AAI64066 D 107.6 107.6 107.6 25 AAI64066 D 107.6 25	~ .			19.7	1032	21			
29 193 19.2 30 169.8 16.7 722 24 AAS61992 31 166.6 16.4 700 24 AAS61993 32 163.8 16.2 481 25 ABZ19574 33 163.4 16.1 300 20 AAZ14552 34 161.2 15.9 1515 21 AAC50415 35 156 15.4 861 24 ABN98824 36 153.4 15.1 737 23 AAS79449 37 147.2 14.5 464 21 AAC46721 38 133.2 13.1 615 22 AAH07116 39 107.6 10.6 1149 23 AAS88030 40 107.6 10.6 3279 23 AAS89559 41 65.6 6.5 432 24 ABN78107 42 65 6.4 487 22 AAI98879 43 65 6.4 487 22 AAI98879 44 53.6 5.3 254 25 ABX31310 45 ANC66448	Ċ.			19.2	387	24			
30 169.8 16.7 722 24 AAS61993 31 166.6 16.4 700 24 AAS61993 32 163.8 16.2 481 25 ABZ19574 33 163.4 16.1 300 20 AAZ14552 34 161.2 15.9 1515 21 AAC50415 35 156 15.4 861 24 ABN98824 36 153.4 15.1 737 23 AAS79449 37 147.2 14.5 464 21 AAC46721 38 133.2 13.1 615 22 AAH07116 39 107.6 10.6 1149 23 AAS88030 40 107.6 10.6 3279 23 AAS89559 41 65.6 6.5 432 24 ABN78107 42 65 6.4 487 22 AAI98879 43 65 6.4 487 22 AAI98879 44 53.6 5.3 254 25 ABX31310 Human GDP-mannose Human GDP-mannose Human GDP-mannose Human GDP-mannose Human GDP-mannose	С	29	195	19.2	387	24			
31 163.8 16.2 481 25 ABZ19574 32 163.8 16.2 481 25 ABZ19574 33 163.4 16.1 300 20 AAZ14552 34 161.2 15.9 1515 21 AAC50415 35 156 15.4 861 24 ABN98824 26 153.4 15.1 737 23 AAS79449 37 147.2 14.5 464 21 AAC46721 38 133.2 13.1 615 22 AAH07116 39 107.6 10.6 1149 23 AAS88030 40 107.6 10.6 3279 23 AAS89559 41 65.6 6.5 432 24 ABN78107 42 65 6.4 487 22 AAI98879 43 65 6.4 487 22 AAI64066 44 53.6 5.3 254 25 ABX31310 Human GDP-mannose Human GPP-mannose Human GPP-mannose Human GPP-mannose Human GPP-mannose		30	169.8	16.7	722				
32 163.8 16.2 401 20 AAZ14552 34 161.2 15.9 1515 21 AAC50415 35 156 15.4 861 24 ABN98824 C 36 153.4 15.1 737 23 AAS79449 37 147.2 14.5 464 21 AAC46721 38 133.2 13.1 615 22 AAH07116 39 107.6 10.6 1149 23 AAS88030 40 107.6 10.6 3279 23 AAS89559 41 65.6 6.5 432 24 ABN78107 42 65 6.4 487 22 AAI98879 43 65 6.4 487 22 AAI64066 44 53.6 5.3 254 25 ABX31310 Human gene express Arabidopsis thalia Arabidopsis thalia DNA encoding novel DNA encoding novel DNA encoding novel Human CDNA clone (DNA encoding novel DNA encoding novel DNA encoding novel Human ORF3054 cDNA Human GDP-mannose Human Bladder rela Human Bladder rela Human GDP-mannose Human GDP-mannose		31	156.6	16.4					
33 163.4 16.1 30 21 AAC50415 Arabidopsis thalia 35 156 15.4 861 24 ABN98824 DNA encoding novel 26 and 27 147.2 14.5 464 21 AAC46721 Zea mays DNA fragm 38 133.2 13.1 615 22 AAH07116 Human cDNA clone (39 107.6 10.6 1149 23 AAS88030 DNA encoding novel 40 107.6 10.6 3279 23 AAS89559 DNA encoding novel 41 65.6 6.5 432 24 ABN78107 Human ORF3054 cDNA 42 65 6.4 487 22 AAI98879 Human excretory re 43 65 6.4 487 22 AAI64066 Human bladder rela 44 53.6 5.3 254 25 ABX31310 Human GDP-mannose Human GDP-mannose		32	163.8	16.2				II was some company	
35		3.3	163.4	16.1					a a
C 36 153.4 15.1 737 23 AAS79449 37 147.2 14.5 464 21 AAC46721 38 133.2 13.1 615 22 AAH07116 39 107.6 10.6 1149 23 AAS88030 40 107.6 10.6 3279 23 AAS89559 41 65.6 6.5 432 24 ABN78107 42 65 6.4 487 22 AAI98879 43 65 6.4 487 22 AAI64066 44 53.6 5.3 254 25 ABX31310 HUMAN CODA encoding novel HUMAN ORF3054 CDNA HUMAN ORF3054 CDNA HUMAN Excretory re HUMAN GDP-mannose HUMAN GDP-mannose HUMAN GDP-mannose HUMAN Secreted pro		34	161.2	15.9			=		
26 153.4 15.1 75.7 25 AAC46721 Zea mays DNA fragm 37 147.2 14.5 464 21 AAC46721 Human cDNA clone (39 107.6 10.6 1149 23 AAS88030 DNA encoding novel 40 107.6 10.6 3279 23 AAS89559 DNA encoding novel 41 65.6 6.5 432 24 ABN78107 Human ORF3054 cDNA 42 65 6.4 487 22 AAI98879 Human excretory re 43 65 6.4 487 22 AAI64066 Human bladder rela 43 53.6 5.3 254 25 ABX31310 Human GDP-mannose Human secreted pro		35	156	15.4					
38 133.2 13.1 615 22 AAH07116 Human cDNA clone (39 107.6 10.6 1149 23 AAS88030 DNA encoding novel 40 107.6 10.6 3279 23 AAS89559 DNA encoding novel 41 65.6 6.5 432 24 ABN78107 Human ORF3054 cDNA 42 65 6.4 487 22 AAI98879 Human excretory re 43 65 6.4 487 22 AAI64066 Human bladder rela 44 53.6 5.3 254 25 ABX31310 Human GDP-mannose 44 53.6 5.3 254 25 ABX31310 Human secreted pro	C	36	153.4						
38 133.2 13.1 013 22 AAS88030 DNA encoding novel 40 107.6 10.6 3279 23 AAS89559 DNA encoding novel 41 65.6 6.5 432 24 ABN78107 Human ORF3054 cDNA 42 65 6.4 487 22 AAI98879 Human excretory re 43 65 6.4 487 22 AAI64066 Human bladder rela 44 53.6 5.3 254 25 ABX31310 Human GDP-mannose 44 53.6 5.3 254 25 ABX31310 Human secreted pro		37	147.2	14.5				Luman cDNA clone	(
40 107.6 10.6 3279 23 AAS89559 DNA encoding novel 41 65.6 6.5 432 24 ABN78107 Human ORF3054 cDNA 42 65 6.4 487 22 AAI98879 Human excretory re 43 65 6.4 487 22 AAI64066 Human bladder rela 44 53.6 5.3 254 25 ABX31310 Human GDP-mannose		38	133.2	13.1				DVAding novo	
40 107.6 10.6 3279 23 ABN78107 Human ORF3054 cDNA 41 65.6 6.5 432 24 ABN78107 Human ORF3054 cDNA 42 65 6.4 487 22 AAI98879 Human excretory re 43 65 6.4 487 22 AAI64066 Human bladder rela 44 53.6 5.3 254 25 ABX31310 Human GDP-mannose		39	107.6	10.6				DITAding notto	1
41 65.6 6.5 432 24 AAI98879 42 65 6.4 487 22 AAI98879 43 65 6.4 487 22 AAI64066 44 53.6 5.3 254 25 ABX31310 Human GDP-mannose 44 53.6 5.3 254 25 ABX31310 Human secreted pro								TO DESCRIPTION OF A COM-	
43 65 6.4 487 22 AAI64066 Human bladder rela 43 65 6.4 487 22 AAI64066 Human GDP-mannose 44 53.6 5.3 254 25 ABX31310 Human secreted pro		41	65.6					TT	
44 53.6 5.3 254 25 ABX31310 Human GDP-mannose		42	. 65					for robbald It	
44 53.6 5.3 244 25 ANCOCA40 Human secreted pro		43	65					TT CDD manage	
45 43 4.2 447 21 AACU6449 Human Secreted pro		44	53.6						
		45	43	4.2	2 44	7 2	1 AACU6449	itumati beeleeda pi	_

```
RESULT 1
AAF86462
     AAF86462 standard; cDNA; 1014 BP.
ID
XX
     AAF86462;
AC
XX
                  (first entry)
     26-JUN-2001
DT'
     Human Acute Neuronal Induced Calcium Binding Protein, ANIC-BP, cDNA.
XX
DE
XX
     Human; cerebroprotective; neuroprotective; vulnerary; vaccine;
     gene therapy; Acute Neuronal Induced Calcium Binding Protein; ANIC-BP;
KW
     stroke; acute head trauma; multiple sclerosis; spinal cord injury; ss.
ΚW
KW
XX
     Homo sapiens.
OS
XX
                      Location/Qualifiers
FH
     Kev
                       1..1014
FT
     CDS
                       /*tag= a
                       /product= "Human Acute Neuronal Induced Calcium Binding
FT
FT
                      Protein, ANIC-BP"
FT
XX
      WO200123552-A1.
 DM.
ЯX
Cr
      05-APR-2001.
 XX.
      18-8MP-2000; 2000WO-EP09132.
 PV
 YX
                      99EP-0118848.
 \mathcal{R}\mathcal{L}
      _A-SEP-1999;
 \mathbf{X}_{\mathcal{X}}
      (MIRE ) MERCK PATENT GMBH.
 ĹΑ
 XX
                    Duecker K;
      Dam Daas I,
 \Sigma
 XX
      WPT: 2001-308142/32.
 ĎŖ
      P-PSDB; AAB82090.
 \mathbb{DR}
 XX
      Novel human acute neuronal induced calcium binding polypeptide, and
      polynucleotides encoding them useful for diagnosing or treating stroke,
 TT.
 PT
      acute head trauma, multiple sclerosis and spinal cord injury -
 \mathbf{p}\mathbf{T}
 XX
      Claim 5; Page 40-41; 45pp; English.
 PS
      The present sequence is the coding sequence for human Acute Neuronals.
 ХX
       Induced Calcium Binding Protein (ANIC-BP). ANIC-BP coding sequence and
 CC
       protein are useful for treating stroke, acute head trauma, multiple
 CC:
       sclerosis and spinal cord injury. ANIC-BP coding sequence and protein
 CC:
       are also useful as vaccines for inducing an immunological response in a
 CC
 CC
       mamaarl.
 CC
 XX
       Sequence 1014 BP: 340 A; 205 C; 209 G; 260 T; 0 other;
  SO
                                                             Length 1014;
                             100.0%; Score 1014; DB 22;
    Query Match
                             100.0%; Pred. No. 3.5e-272;
    Best Local Similarity
                                                                                    0;
    Matches 1014; Conservative 0; Mismatches
                                                                           Gaps
                                                       0; Indels
```

Qy	1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Db	3. ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGTATA
Qy	61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC
Db	TO A TICA CITICA A COA A TICA A ACA A ATTICTGTGTGTACAAACGAGAAAGAA 180
Qy	121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGATTTTTTTT
Db	
Qy	181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Db	181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGGGTGGGT
Qy	241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
Db	241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAA
Qy	301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Db	301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGGAGTCCTACTGTGGGAGTCCTACTGTGGGAGTCCTACTGTGGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGGAGTCCTACTGTGAGTCCTACTGTGAGTACTGTGAGTCCTACTGTGAGTCCTACTGTGAGTCCTACTGTGAGTCCTACTGTGAGTCCTACTGTGAGTCCTACTGTGAGTCCTACTGTGAGTCCTACTGTGAGTCCTACTGTGAGTCCTACTGTGAGTCTACTGTGAGTCCTACTGTACTGTGAGTCCTACTGTACTACTGTACTACTGTACTACTGTACTACTGTACTACTACTGTACTACTACTACTACTACTACTACTACTACTACTACTACT
Qy	361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
Db dc	TO TREE TO THE TREE TREE TO THE TREE TREE TREE TREE TREE TREE TREE
Qy Do.	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGTATGT
Cy .	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
Db _i	161 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTAACGTACGTGGAGTTGTCMAGATTTTAAGTACGTAGATTTTAAGTACGTGGAGTTGTCMAGATTTTAAGTACGTAGATTTTAAGTACGTAGATTTTAAGTACGTAGATTTTAAGTACGTAGATTTTAAGTACGTAGATTTTAAGTACGTAGATTTTAAGTAACGTAGATTTTAAGTAACGTAGATTTTAAGTAACGTAGATTTTAAGTAACGTAGATTTAAGTAACGTAGATTTTAAGTAACGTAGATTTTAAGTAACGTAGATTTAAGTAACGTAGATTTTAAGTAACGTAGATTTAAGTAACGTAGATTTAAGTAACGTAGATTTAAGTAACGTAGATTTAAGTAACGTAGATTTAAGTAACGTAGATTTAAGTAACGTAGATTTAAGTAACGTAGATTTAAGTAACGTAGATTTAAGTAACGTAGATTAAGTAACGTAGATTAAGTAACATTAAGTAACATAAGTAACATAAGTAACATAAGTAACATAAGTAACATAAGTAACATAAGTAACATAAAGTAACATAAAGTAACATAAAGTAACATAAAGTAACATAAAAAAAA
QУ	541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600.
Гb	
Qу	691 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
DB	
Qγ	661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
Do "	651 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
Qy	721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
Db	721 AACTTTGCCATCATGACAAAATTTTTTTTTTTTTTTTTT
Qv	781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAACGTGTTTT 840
Db	
Qу	341 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900

```
841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
Db
         901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
             Qу
         901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
Db
         961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
             QУ
         961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Db
RESULT 2
AAI58234
     AAI58234 standard; cDNA; 1421 BP.
TD
XX
     AAI58234;
AC'
XX
     22-OCT-2001 (first entry)
DT
ХХ
     Human polynucleotide SEQ ID NO 437.
DΕ
     Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX
     peripheral nervous system; neuropathy; central nervous system; CNS;
ΧW
     Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XW
     amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
\mathbb{Z}\mathbb{W}
     chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
MS
KH
     leukaemia; ss.
20!
XY_{\bullet}
     Momo sapiens.
 OS
             4 4
 XX
     4:0200153312-A1.
 PM
 26-JUL-2001.
 DD.
 XX
      25-DEC-2000; 2000WO-US34263.
 PF
 XX
      21-JAN-2000; 2000US-0488725.
 PR
      25-APR-2000; 2000US-0552317.
 \mathbb{R}
      09-JUL-2000; 2000US-0598042.
 \mathbf{p}\mathbf{p}
      19-JUL-2000; 2000US-0620312.
 PR
      03-AUG-2000; 2000US-0653450.
 DΚ
      14-SEP-2000; 2000US-0662191.
 DR
      19-OCT-2000; 2000US-0693036.
 PR
      29-NOV-2000; 2000US-0727344.
 PR
 XX
      (HYSE-) HYSEQ INC.
 PΑ
 XX
                                                                  Wang D;
               Liu C, Asundi V, Chen R, Ma Y,
                                                          Ren F,
                                                 Qian XB,
      Tang YT,
 P.I
                       Wehrman T, Xu C, Xue AJ, Yang Y,
                                                          Zhang J;
              Wang Z,
      Wang J,
 PΙ
                Zhou P, Goodrich R, Drmanac RT;
 ΣŢ
      Zhao QA,
 XX
      WPI; 2001-442253/47.
 DR
      P-PSDB; AAM39078.
 DR
  XX
      Novel nucleic acids and polypeptides, useful for treating disorders
  PT
      such as central nervous system injuries -
  PT
```

```
XX
   Claim 1; SEQ ID NO 437; 10078pp; English.
PS
XX
   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC
   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC
    immunosuppressant and cytostatic activity. The polynucleotides are useful
CC
    in gene therapy. A composition containing a polypeptide or polynucleotide
CC
    of the invention may be used to treat diseases of the peripheral nervous
CC
    system, such as peripheral nervous injuries, peripheral neuropathy and
CC
    localised neuropathies and central nervous system diseases, such as
CC
   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC
    lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC
    utilisation of the activities such as: Immune system suppression,
CC
    Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC
    and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC
    assays for receptor activity, arthritis and inflammation, leukaemias and
CC
    C.N.S disorders.
CC
    Note: The sequence data for this patent did not form part of the printed
CC
    specification.
CC
XX
    Sequence 1421 BP; 469 A; 284 C; 306 G; 362 T; 0 other;
SÓ
                          Score 1014; DB 22; Length 1421;
                    100.0%;
 Query Match
                          Pred. No. 4e-272;
                    100.0%;
  Best Local Similarity
                            Mismatches
                                          Indels
                                                    Gaps...
  Matches 1014; Conservative
                         0;
         1 ATGAAAAAAATGCCTTTGTTAGTAAATCACAGAAAAATCCAGCAGAAATTGTGAAAATC 60
Qy:
           217 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 276
Db
        够
           כלם
        \mathfrak{D}\lambda
           Db
        381 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
\Omega \lambda
           297 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 456
Db
        241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
QΥ
           457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 516
Db
        301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Qy
           517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576
Db
        361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
Qγ
           577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636
Db
        421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
QУ
```

637 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 696

Db

```
481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
                        Qу
                       TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 756
Db
                 \pm 4.1 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC \pm 6.00
                        QУ
                       TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 816
Db
                 601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
                         Qy
                 817 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 876
Db
                  561 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
                         QУ
                  877 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 936
Db
                  721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
                         Qу
                  937 AACTITGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 996
 O,O
                  781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
                         Qy
                  997 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 1056
 Db
                   (A1 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 90.0)
                         1
                  .057 CTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1116
 200
                  DOL STOATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960%
                          THE THE PROPERTY OF THE PROPER
                 1117 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1176
 Di,
                   961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
                          ÕΆ
                 1177 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1230
  Db
  RESULT 3
  AAA27332
          AAA27332 standard; cDNA; 1344 BP.
  ID
  XX
  FC
           MAA27332;
  XX
                                  (first entry)
           40-AUG-2000
  DT
  XX
           Human calcium binding protein hCBP gene.
  DE
   XX
           Hur,an; calcium binding protein; cancer; inflammation; CBP;
   KW
           reproductive disorder; autoimmune disorder; developmental disorder;
   KW
           seizure disorder; immune disorder; infection; ss.
   KW
   XX
           Homo sapiens.
   OS
   XX
                                       Location/Qualifiers
   FH
           Kev
                                       124..1134
           CDS
   FT
                                        /*tag=
   FT
```

```
/product= "calcium binding protein"
FT
XX
    WO20C029580-A1.
PN
XX
    25-MAY-2000.
PD
XX
                 99WO: US27027.
    12-NOV-1999;
PF
XX
                 98US-0190965.
    13-NOV-1998;
PR
XX
    (INCY-) INCYTE PHARM INC.
PA
XX
                      Corley NC,
                                  Gorgone GA;
    Tang YT, Guegler KJ,
PΙ
XX
    WPI; 2000-387793/33.
DR
    P-PSDB; AAY94247.
DR
    Human hCBP protein, and the nucleic acid encoding it, useful for e.g.
XX
    diagnosis, prevention and treatment of cancers, immune, developmental
PT
PT
    or reproductive disorders -
PT
XX
    Claim 9; Fig 1; 72pp; English.
PS
    The present sequence is the human calcium binding protein hCBP gene. It
XX
CC
    was obtained by screening a coronary artery smooth muscle cDNA library,
CC
    from which five overlapping nucleic acids were isolated and
    sequenced, and then expressed to give the protein. The protein and the
CC
    gene encoding it are useful for the diagnosis and treatment of the
CC
    following types of disorder: cancers (such as adenocarcinomas);
CC
55.
    reproductive disorders (such as infertility, ovulatory defects);
63
     endometriosis, disruptions of the oestrus and menstrual cycles;
     polycystic ovary syndrome and ovarian hyperstimulation), autoimmune
CC
     disorders (such as benign prostatic hyperplasia and prostatitis),
CC
     developmental disorders (such as Cushing's syndrome, muscular dystrophy
CC
     and gonadal dysgenesis), hereditary neuropathies, seizure disorders,
 CC.
 CO
     immune disorders (such as AIDS, allergies, anaemia, asthma,
     atherosclerosis, cholecystitis, Crohn's disease, diabetes, Graves'
 CC
 CC
     disease, multiple sclerosis, psoriasis, rheumatoid arthritis,
     scleroderma, Sjogren's syndrome and ulcerative colitis), and viral,
 CC
 CC
     pacterial, fungal, parasitic, protozoal and helminthic infections.
 CC
 XX
     Sequence 1344 BP; 450 A; 261 C; 280 G; 353 T; 0 other;
 SO
                              Score 1010.8; DB 21; Length 1344;
                       99.7%;
   Query Match
                              Pred. No. 3.1e-271;
                       99.8%;
   Best Local Similarity
                                                                   0;
                                                        0; Gaps
                                               Indels
                             0; Mismatches
   Matches 1012; Conservative
           1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
             Qу
         1.24 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 183
 Db
          QΥ
          Db
          Qy
```

Db	:	244	GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA
Qy			CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Db			•
Qy			ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGGAAAAAAGATGTGACCCAGATA 300
Dр			
QУ			TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
מת		,	
ÖÄ			GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
Db			
Qy D£	v.		CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
			540
DP ÖM "			TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTGGGTGTGTGATATTGATATTGCT 663
45			PCAGATGCCTTTGCTACTTTCAAGGAFTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 500
υ. Σ	0		*CAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 72354 *CCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 72354
QZ Sib		401 734	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660 ;
7			GENERAL CHARTETTANI GCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
Q}′ Db ≐			AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTTTAGAGGTGCTTAGAGGTGCTTAGAGGTGCTAGGGGGAGCTGATCCTGGACCGTCAC 843-
			THE CONTRACTOR OF A CITATIATICA COA AGCCGGAGAACCTGAAACTCATGATGAAC 780:
ĞΥ			
Db			4 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGATGATA
Qy			1 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
Dr			
ĞУ	*		E GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 9004
Db.			the state of the s
QУ	•		1 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
Db			•
QУ		:	GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Db		108	34 GAGAAGAACIACIIGAIII

おおから 一大大大大

```
RESULT 4
AAH15879
     AAH15879 standard; cDNA; 2002 BP.
ID
XX
     AAH15879;
AC
ХХ
                   (first entry)
     26-JUN-2001
DT
XX.
     Human cDNA sequence SEQ ID NO:14407.
DE
     Human: primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
ΚW
XX
     Homo sapiens.
os
XX
      EP1074617-A2.
PN
XX
      (F)-FEB-2001..
PD
XX
      20-JUL-2000; 2000EP-0116126.
 PF.
 YX
                     99JP-0248036.
      %9-JUL-1999;
 FR
                     99JP-0300253.
      -AUG-1999;
 bS
      11-0AN-2000; 2000JP-0118776.
 1-13
      M-MAY-2000; 2000JP-0183767.
 £R
      19-JUN 2000, 2000JP-0241899.
 PR.
 XX
      (HETI-) BELIX RES INST.
 :: <sub>2</sub>.
                                                                      x >
 4525
      ියෙ ී, Isogai T, Nishikawa T, Hayashi K, Saito K,
                                                              Yamamoto 👵 👯
      Shii S. Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 N_{\rm eff}
      WPT; 2001-318749/34.
 QP:
      Primer sets for synthesizing polynucleotides, particularly the 5602
 XX
      full-length cDNAs defined in the specification, and for the detection
 \mathbf{T}^{\mathbf{r}}
      and/or diagnosis of the abnormality of the proteins encoded by the
 PT
 PΤ
       full-length cDNAs -
 PT
       Claim 8; SEQ ID 14407; 2537pp + CD ROM; English.
 XX
 PS
       The present invention describes primer sets for synthesising 5602
 XX
       full-length cDNAs defined in the specification. Where a primer set
  CC
      comprises: (a) an oligo-dT primer and an oligonucleotide complementary
  CC
       to the complementary strand of a polynucleotide which comprises one of
  CC
       the 5602 nucleotide sequences defined in the specification, where the
  CC
       oligonucleotide comprises at least 15 nucleotides; or (b) a combination
  CC
       of an oligonucleotide comprising a sequence complementary to the
  CC
       complementary strand of a polynucleotide which comprises a 5'-end
  CC
       sequence and an oligonucleotide comprising a sequence complementary to a
  CC
  CC
       polynucleotide which comprises a 3'-end sequence, where the
       oligenucleotide comprises at least 15 nucleotides and the combination of
  CC
       the 5'-end sequence/3'-end sequence is selected from those defined in
  CC
       the specification. The primer sets can be used in antisense therapy and
  CC
       in gene therapy. The primers are useful for synthesising polynucleotides,
  CC
       particularly full-length cDNAs. The primers are also useful for the
  CC
```

CC

the full-length cDNAs. The primers allow obtaining of the full-length CDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 CC represent oligonucleotides, all of which are used in the exemplification CC CC of the present invention. CC XXSequence 2002 BP; 594 A; 418 C; 463 G; 527 T; 0 other; SQ Length 2002; Score 992.8; DB 22; 97.9%; Query Match Pred. No. 3.8e-266; 99.8%; Best Local Similarity 0; Indels. Gaps Mismatches. 994; Conservative 0: Matches 19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAAGACAATTTGGCC 78 Ωy 1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60 Dр 79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138 Qÿ 51 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120 Db 139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAGAACCCCCCAACAGAAGCAGTG 198 ΰУ CAAGCAATGAAAGAAATTCTGTGTGTGCTACAAACGAGAAGAACCCCCCAACAGAAGCAGTG 180 Jb 199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258 131 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240 1.15 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318 4 QY CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300, Ωħ DES AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 373 12.7 301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360 Db 379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438 ΩV 361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420 Db 439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498 () y 421 AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480 ŢĴŹ9 49.9 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558 Q_T^{ω} 481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540 Db 559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618 ुर: 541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600 Db 678 GACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678 ÇΥ 601 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660

detection and/or diagnosis of the abnormality of the proteins encoded by

CC

CC

Db

```
679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
QУ
            661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720
Db
         739 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
. Qу.
            721 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 780
Db
         799 CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAA 858
Qу
            781 CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAA 840
Db
         359 ACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 918
QУ
            341 ACACAGCCTATTGTGGAGATGCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 900
Db
         919 AGCTTCCAAAAAGAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 978
 Qу
            901. AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 960
 ď
         979 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
 Qy
            961 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 996
 Db.
 RESULET 5
 AME05471
    #EMO5471 standard; cDNA; 322 BP.
 XX
 AC
     AAH05471;
 XX
     36-JUN-2001 (first entry)
 DT
 ХX
     Human cDNA clone (5'-primer) SEQ ID NO:2306.
 DE
 XX
     Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss:
 ХW
 XX
     Momo sapiens.
 OS
 XX
 PIJ
     EP1074617-A2.
 XX
     07-FFB-2001.
 PD
 XX
     ЭЗ-ЛЛL-2000; 2000EP-0116126.
 PF
 XX
     29-JUL-1999;
                  99JP-0248036.
 PR
                  99JP-0300253.
     27-AUG-1999;
 PR
     11-JAN-2000; 2000JP-0118776.
 PR
     02-MAY-2000; 2000JP-0183767.
 PR
     09-JUN-2000; 2000JP-0241899.
 PR
 XX
      (HELI-) HELIX RES INST.
 DΔ
 XX
     Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K,
                                                   Yamamoto J;
 PI
     Ishii S, Sugiyama T, Wakamatsu A, Nagai K,
 PI
```

XX

DR WPI; 2001-318749/34.

XX PΤ

PT

PT

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

PTXXPS

Claim 1; SEQ ID 2306; 2537pp + CD ROM; English.

XX CC

CC

CC

CC CC

CCCC

CC

CC

CC

CC

CC

CC CC

CC.

CC

CC

W.

TO:

CC

17

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

@C XXSQ .

Sequence 822 BP; 268 A; 164 C; 171 G; 216 T; 3 other;

Query Match 76.0%; Score 770.6; DB 22; Length 822; Eest Local Similarity 98.5%; Pred. No. 2.1e-204; Matches 798; Conservative 0; Mismatches Indels 2; Gaps : .

19 TITAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78 Dp-1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60 79 ATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAGTGTCTAAATCACTG 138 ĎΜ 761. ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120 ·Dib US9 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAGAACCCCCAACAGAAGCAGTG 198 Qу 121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAGAACCCCCAACAGAAGCAGTG 180 Db Qу 199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258 181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240 Dio 259 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318 Qy

Db 241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300

```
319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
          Qу
       301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360
Db
       379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
           ÕÃ
       361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420
Db
       439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498
           Qу
          AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480
Db
       429 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
           Qγ
        481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540
Db
        559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
           QУ
        541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600
Lb
        619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
           Qу
        501 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660
Db
        CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
           170
           CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTT'IGCCATCATGACA 729:1
Db
        70% AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
           AAGTATATCAGCAAGCCGGAGAACCTG-AACTCATGATGAACCTNCTTCGGGAT-AAAGT 778
Dir
        759 CCCAACATCCAGTTTGAAGCCTTTCATGTT 828
QУ
           779 CCCAACATCCAGTTTGAGCCTTCTGGTTTT 808
Db
RESULT 5
AAX39817
    AAX39817 standard; DNA; 831 BP.
T.F)
XX
    AAX39817;
AC
XX
    02-JUL-1999 (first entry)
DΤ
XX
    Castric cancer associated gene.
 DE
 XX
    Cancer associated antigen; diagnosis; research; treatment; human;
 KW
    breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
 YV
     prostate cancer; ss.
 K\mathbb{N}
 XХ
     Homo sapiens.
 0S
 XX
     W09904265-A2.
 PN
 XX
     28-JAN-1999.
```

PD

```
XX
                98WO-US14679.
    15-JUL-1998;
PF
XX
                98US-0102322.
    22-JUN-1998;
PR
                97US-0896164.
    17-JUL-1997;
PR
                97US-0061599.
    10-CCT-1997;
PR
                97US-0061765.
    10-OCT-1997;
PR
                97US-0948705.
    10 OCT-1997;
PR
                97GB-0021697.
    11-OCT-1997;
PR
XX
    (LUDW-) LUDWIG INST CANCER RES.
PA
XX
                          O'Hare M, Obata Y, Old LJ;
    Chen Y, Gout I, Gure A,
Τσ
    Pfreundschuh M, Sahin U, Scanlan MJ,
                                     Stockert E;
PΙ
    Tureci 0;
IJΤ
XX
    WPI; 1999-132448/11.
DR
    New isolated cancer associated nucleic acids and polypeptides -
XX
    isolated using sera from cancer patients, used to develop products
יוֻיםַ
PŢ
    for the diagnosis, monitoring or treatment of cancers
PT
XX
    Chaim 67; Page 558-559; 787pp; English.
PS.
    The invention relates to a method for diagnosing a disorder characterised
XX
    by expression of a human cancer associated antigen precursor coded for by
CC
    a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CII
    biological sample isolated from a subject with an agent that specifically
00
    binds to the NAM, an expression product or a fragment of an expression
00
OC.
    product complexed with an HLA molecule; and (b) determining the
    interaction between the agent and the NAM or the expression product as a
CC
CC.
    determination of the disorder. The products and methods can be used in
 CC.
    the diagnosis, monitoring, research, or treatment of conditions
     characterised by the expression of various cancer associated antigens.
     The invention provides nucleic acid sequences and encoded polypeptides
 CC.
     which are cancer associated antigen precursors expressed in human breast
 CC
     cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
 CC
     lung dancer.
 CC
 XX
     Sequence 831 BP; 285 A; 165 C; 167 G; 209 T; 5 other;
 SQ
                             Score 684.6; DB 20; Length 831;
                      67.5%;
  Query Match
                      96.1%; Pred. No. 1.9e-180;
  Best Tocal Similarity
                                              Indels
                                                          Gaps
                            0; Mismatches
                                          23;
  Matches 764; Conservative
           1. ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
             ŨУ
          37 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 96
 Db
          QY
          Db
         QУ
         Db
```

```
181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
           QУ
       217 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 276
Db
       241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
          Qy
       277 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 336
ηb
       301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
           Qу
       337 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 396
Db
        361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
           ΩУ
        397 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 456
Db
        421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
           QУ
        457 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 516
Db
        481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
           QУ
        517 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 576
Do
        54% TCAGATGCCTTTGCTACTTTCAA-GGATTTACTAACCAGACATAAAGTGTTGGTAGC-AG 598
          ĜĀ
        CAGATGCCTTTGCTACTTTCAAGGGATTTACTAACCAGACATAAAGTGTTGGTAGCAAG 636
Db
        ACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTG 6589
           ΟУ
        ACTTCTTAGAACAAAATTACGACACTANTTTTGAAGACTATGAGAAATTGCTTCAGTCTG 696:7
 550 AG-AATTATGTTACTAAGAGACAGTCTTTAAAG-CTGCTAGGGGAGCTGATCCTGGACCG 7165 8
           QΥ
        597 AGAAATTATGTTACCAAGAGACAGTCCTTAAAGCCTGCTAAGGGAACTGATTCTGGACCG 756.
 Db
        717 TCACAACTTTGCCATC-ATGACAAAGTATATCAGCAAGCC---GGAGAACCTGAAACTCA 772
                                            | | | | | |
 \delta \lambda
                            757 TCANAACTTTGCCATCAANGCAAAAGTTTATCAACAAGCCNGGGGAAACCGGAAACNCAA 816
 Db:
         773 TGATGAACCTCCTTC 787
 QУ
             817 GGAGGAACCTCCTTC 831
 Do
 PESULT 7
 AAI60020/C
     AAI60020 standard; cDNA; 1191 BP.
 ED
 XX
     AAI60020;
 AC
 XX
               (first entry)
     22-OCT-2001
 DT
 XX
     Human polynucleotide SEQ ID NO 4009.
 DE
     Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX
     peripheral nervous system; neuropathy; central nervous system; CNS;
  ΚW
  KW
```

```
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
     amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW
     chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
ΚW
KW
     leukaemia; ss.
KW
XX
     Homo sapiens.
OS
ХX
     WO200153312-A1.
PN
XX
     26-JUL-2001.
PD
XX
     26-DEC-2000; 2000WO-US34263.
PF
XX
     21-JAN-2000; 2000US-0488725.
PR
     25-APR-2000; 2000US-0552317.
PR
     09-JUL-2000; 2000US-0598042.
PR
     19-JUL-2000; 2000US-0620312.
PR
     03-AUG-2000; 2000US-0653450.
PR
      14-SEP-2000; 2000US-0662191.
ÐΒ
      19-OCT-2000; 2000US-0693036.
₽R.
      29-NOV-2000; 2000US-0727344.
PR -
XX
      (HYSE-) HYSEQ INC.
 PA
 XX
                                                               Ren F,
      Ying YE, Liu C, Asundi V. Chen R, Ma Y,
                                                    Qian XB,
 \mathfrak{I}\mathfrak{S}
      Hawy of Wang Z, Wehrman T, Ku C, Xue AJ,
                                                     Yang Y,
                                                               Zhang J;
 TI
      Time QA, Zhou P, Goodrich R, Drmanac PT;
 177
 Z.,.
      NPI; 2001-442253/47.
 15.A.
      P-PSDB; AAM40864.
 2)6'
      Movel nucleic acids and polypeptides, useful for treating disorders,
 110
 22
      ruch as central nervous system injuries --
 54
 XX.
      Claim 1; SEQ ID NO 4009; 10078pp; English.
 PS
      The invention relates to human nucleic acids (AAI57798-AAI61369) and
 XX
      the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC
       immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC
       in gene therapy. A composition containing a polypeptide or polynucleotide
 CO
       of the invention may be used to treat diseases of the peripheral nervous
 CC
       system, such as peripheral nervous injuries, peripheral neuropathy and
 CC
       localised neuropathies and central nervous system diseases, such as
 CC
       Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC
       lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
  CC
       atilisation of the activities such as: Immune system suppression,
  CC
       Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
  CC
       and thrombolytic activity, cancer diagnosis and therapy, drug screening,
  ĊC
       assays for receptor activity, arthritis and inflammation, leukaemias and
  CC
  CC
       Note: The sequence data for this patent did not form part of the printed
  CC
  CC
       specification.
  CC
       Sequence 1191 BP; 348 A; 261 C; 236 G; 346 T; 0 other;
  XX
  SQ:
                             67.5%; Score 684.4; DB 22; Length 1191;
    Query Match
                                     Pred. No. 2.6e-180;
                             99.9%;
    Best Local Similarity
```

```
0;
                                                   Gaps
                                                0:
                                        Indels
                                     1;
                           Mismatches
 Matches 685; Conservative
                        0:
       329 GCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCC 388
          QУ
      1189 GCACTCGAAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCC 1130
Db
       389 TCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTA 448
          Qу
      1125 TCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTA 1070
Db
       449 TTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTA 508
          Qу
       1069 TTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTA 1010
Db
       509 AGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATT 568
           QΥ
       1009 AGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATT 950
Db
        569 TACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTT 628
           QV
        949 TACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTT 890
Db
        529 TTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAA 688
           Q.
          TTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGACACAGTCTTTAA 830
DD
        AND AGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCA 74822
           \mathcal{Q}_{i}
           AGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCA 77.0
Db
        745 CCAACCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCC 308.
           QΨ
        759 GCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCC 710
 T Do
        309 AGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAAACACAGCCTA 868
           \Omega
        709 AGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAAACACAGCCTA 650
 \Gamma
        869 TTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAA 928
            649 TTGTGGAGATCCTGTTAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAA 590
 Db
         929 AAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCC 988
            Ο.A.
        539 AAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCC 530
 Db
         989 GAGACTTGAAGAAAACGGCCCCTTGA 1014
 QΥ
            111111111111111111111111111111
         529 GAGACTTGAAGAAAACGGCCCCTTGA 504
 Db
```

RESULT 8

AAC91772 ID AAC91772 standard; cDNA; 1026 BP.

XX

AC AAC91772;

XX

```
27-MAR-2001 (first entry)
DТ
    Human ANIC-BP (acute neuronal induced calcium-binding protein) cDNA.
XX
DE
     Human; acute neuronal induced calcium-binding protein; ANIC-BP;
XX
     Mo25 homologue; HymA homologue; drug screening; stroke;
ΚW
     acute head trauma; multiple sclerosis; spinal cord injury; vaccine;
KW
ΚW
     cerebroprotective; neuroprotective; ss.
KW
XX
     Homo sapiens.
os
XX
     WO200078947-A1.
PN
XX
     28-DEC-2000.
PD
XX
     14-JUN-2000; 2000WO-EP05457.
PF
XX
                    99EP-0112024.
     22-JUN-1999;
 ΡR
XX
      (MERE ) MERCK PATENT GMBH.
 PΑ
 XX
     Den Daas I, Fischer V, Seyfried C,
                                           Von Melchner L;
 PΙ
 XX
      WPI; 2001-102721/11...
 DR
      P-PSDB; AAB48970.
 DR.
      Movel acute neuronal induced calcium binding protein, useful for
 XX
      treating acute head trauma, stroke, multiple sclerosis and spinal
 21
 \geq T
 327
      irjury
 XX
      Claim 5; Page 35-36; 50pp; English.
 38
      The invention relates to human acute neuronal induced calcium-binding
 XX
      protein (ANIC-BP) and to nucleic acid encoding it. The invention
 COS
      also relates to expression systems and recombinant host cells comprising
 CC
      ANIC-BP DNA, the recombinant production of ANIC-BP, antibodies specific
 CC.
      for ANIC-BP, fusion proteins comprising ANIC-BP and an immunoglobulin
 CC
      Fc region, and methods of screening for modulators of ANIC-BP function.
 CC
      ANIC-BP has homology and structural similarity to HymA and Mo25 proteins.
 CC
      ANIC-BP proteins and nucleotides are useful for treating stroke and
  CC
      acute head trauma, multiple sclerosis and spinal cord injury. ANIC-BP
  CC
       proteins are useful in screening assays, for identifying membrane bound
  CC
      or soluble receptors, and also in vaccines. ANIC-BP nucleotides are
  CC
       useful as diagnostic reagents, as tools for tissue expression studies,
  CG
      for chromosome localisation studies, as genetic vaccines, and in
  CC
       the generation of transgenic animals. The present sequence represents
  CC
  CC
      cDNA encoding human ANIC-BP.
  CC
       Sequence 1026 BP; 359 A; 199 C; 203 G; 265 T; 0 other;
  XX
  SQ
                                                          Length 1026;
                                    Score 582.6; DB 22;
                            57.5%;
    Query Match
                                    Pred. No. 5.5e-152;
                            74.7%;
    Best Local Similarity
                                                         Indels
                                      Mismatches 244;
    Matches 748; Conservative
                                   0;
             18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
                Qy '
             12 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 71
  DD
```

Qy	78 CATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
Db	
QУ	129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA
D.b	189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
Qy	189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTATATATA
Dp	249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAA 308
QУ	249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAAATAAAAAAAA
Db	309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
Õλ	
рþ	TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 371
Qy-	369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
nb .	372 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 434
Qу "	429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
dG.	ATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 491
$\sqrt{s} \sqrt{g}$	TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548-
ζ(ξ)	ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTGAGATGC 334
E.Z.	CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 6084
Db	552 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGTAGAGTGCAGTAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTA
Qy	509 ACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
.Dp.	
QΥ	569 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728.
DL	•
QУ	CATCATGACAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
Db	
ØУ	789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAG 848
dq	AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 851
Qy	849 TCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
Db	

```
909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
                                                                111111111
              QУ
          912 GTTCCTCAGCAAGTTTCAGAACGACAGGACGAGAGGATGAGCAGTTTAACGACGAGAAGAC 971
Db
          969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
              FIL II - ITHII HIRIT I II - IHII I I I I I I I I
ÇУ
          972 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1012
Db
RESULT 9
ABK13127
     ABK13127 standard; cDNA; 3281 BP.
ID
XX
     ABK13127;
AC
XX
     09-APR-2002 (first entry)
DT
     Human secretory polynucleotide (sptm) cDNA (481257.3).
XX
DΕ
     Signal peptide; transmembrane domain; human; sptm; ss; gene;
 XX
      481257.3; antiarteriosclerotic; antiatherosclerotic; antipsoriatic;
 KW
      autiinflammatory; cytostatic; anti-HIV; antiallergic; antidiabetic;
 KW
      nephrotropic; antigout; antithyroid; hepatotropic; neuroprotective;
 KW
      osseopathic; antirheumatic; antiarthritic; dermatological; cancer;
 4\%I
      inmunosuppressive; antiulcer; ophthalmological; vulnerary; gout;
 KW
      anticonvulsant; cerebroprotective; nootropic; antiparkinsonian;
 KW
      varueide; antibacterial; cell proliferative disorder; arteriosclerosis;
 I:W
      atherosclerosis; psoriasis; immune system disorder; inflammation;
 ICW
      acquired immunodeficiency syndrome; AIDS; Addison's disease;
 W
      adult respiratory distress syndrome; allergy; cirrhosis; osteoporosis;
 WM
      diabetes mellitus; Graves' disease; multiple sclerosis; osteoarthritis;
 KN
      rheumatoid arthritis; systemic lupus erythematosus; ulcerative colitis;
 RW
      haematopoietic cancer; neurological disorder; stroke; epilepsy;
 2.77
      Huntington's disease; Parkinson's disease; meningitis; prion disease;
 KW
      kuru; Creutzfeldt-Jakob disease; cerebral palsy; myasthaenia gravis;
 KŶĬ
 Kii
      diabetic neuropathy; Alzheimer's disease.
 KW
 YX
       Homo sapiens.
  οĎ
  XX
       WO200111032-A1.
  PΝ
  ΧX
       15-FEB-2001.
  PD
  XX
       01-JUN-2000; 2000WO-US15246.
  \mathbb{P}\mathbb{F}
 XX
                      99US-147500P.
       05-AUG-1999;
  PR
                      99US-147501P.
       05-AUG-1999;
  PR
  XX
       (INCY-) INCYTE GENOMICS INC.
  PA
                                 Russo FD, Spiro PA,
Cohen HJ, Rosen BH,
  XX
                                                        Banville SC;
       Hodgson DM, Lincoln SE,
                                                        Chalup MS, Hillman JL;
  PI
       Bratcher SR, Dufour GE,
       Jones AL, Yu JY, Greenawalt LB, Panzer SR,
  PI
                                                       Roseberry AM;
  PI
       Wright RJ, Daniels SE;
  PI
  XX
       WPI; 2002-147236/19.
  DR
```

XX

Novel secretory polynucleotide (sptm) and polypeptides encoded by sptm, useful for diagnosing and treating disorders or diseases associated with cell signaling e.g., allergy, psoriasis, Grave's disease, epilepsy

Claim 1; Page 192-193; 198pp; English.

PT

PT

PT PT XX

PS

CC

CC.

CC

CC

CC

CC

Qu.

cá

CC

CC

aa aa

CC.

 $\mathcal{E}_{\mathbf{Q}}$

Chary Match

This invention relates to novel cDNA molecules encoding isolated secretory polynucleotides (sptm) with similarity to signal peptide (SP) or transmembrane domain (TM) consensus sequences. The polynucleotide sequences of the invention are useful for producing sptm protein by recombinant techniques, the protein may be used to generate anti-sptm antibodies which may be used to analyse protein expression levels in different tissues. The sptm molecules are useful for diagnostic and therapeutic purposes e.g., to diagnose or treat a condition associated with cell signaling such as a cell proliferative disorders (e.g., arteriosclerosis, atherosclerosis, psoriasis, cancers), immune system disorders (e.g., inflammation, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, cirrhosis, diabetes mellitus, gout, Graves' disease, multiple sclerosis, osteoarthritis, osteoporosis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative colitis and haematopoietic cancer), a neurological disorder (e.g., stroke, epilepsy, Huntington's disease, Parkinson's disease, meningitis, prion diseases including kuru, dreutzfeldt-Jakob disease, cerebral palsy, myasthenia gravis, diabetic nouropathy and Alzheimer's disease). Sptm sequences can be used to were detect the presence of or quantifying the amount of sptm-related polynucleotide in a sample. The sptm polynucleotide is used to design probes useful in diagnostic assays carried out to detect or confirm conditions, disorders, or diseases associated with abnormal levels of Sptm expression. Sptm, its fragments or oligonucleotides derived from optm may be used as primers in amplification steps prior to hybridisation. The present sequence represents the human sptm (481257.3) cDNA pequence of the invention,

Sequence 3281 BP; 1014 A; 601 C; 676 G; 990 T; 0 other;

57.5%;

Pred. No. 9.3e-152; 74.7%; Best Local Similarity Indels ·Gaps : Mismatches 244; Matches 748; Conservative 0; 13 GTTTAGTAAATCACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77 Qγ 101 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 160 Db 78 CATTTTGGAAAAGCAAGAC------AAAAAGACAGACAAGGCTTCAGAAGAAGIGTC 128 Ω¥ 11411 1114 111411 4411414111 161 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAAGGCTACAGAAGAAGATTTC 220 Db $C\lambda$ 111221 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 280 Dυ 189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248 Qу]}}}| 281 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 340 Db

Score 582.6; DB 24;

Length 3281;

QУ	249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
Db	341 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 400
Qy Db	369 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
Qy Db	369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
Qy Db	429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
Qy Db	489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
Qy Db	949: CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
Qy	609 BCAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
Cy SW	CACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728 1
2) Db	CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
Çy Do	789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
Qy,	849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
Qy Db	909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
Dp Öλ	969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009

c .

RESULT 10 ABV22987

ID ABV22937 standard; cDNA; 3849 BP.

```
ABV22987;
AC
XX
     13-SEP-2002 (first entry)
DT
XX
     Human prostate expression marker cDNA 22978.
DΕ
XX
     Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
ΚW
KW
     pharmacogenomic marker; gene; ss.
XX
     Homo sapiens.
OS
XX
PN
     WO200160860-A2.
XX
     23-AUG-2001.
PD
XX
     20-FEB-2001; 2001WO-US05171.
PF
XX
     17-FEB-2000; 2000US-183319P.
₽R
     16-MAR-2000; 2000US-189862P.
PR
     25-MAY-2000; 2000US-207454P.
PR
     09-JUN-2000; 2000US-211314P.
₽R
     18-JUL-2000; 2000US-219007P.
PR
     13-DEC-2000; 2000US-255281P.
\mathfrak{RR} .
XX
     (MII-L-) MILLENNIUM PREDICTIVE MEDICINE INC.
\mathbb{P}\mathbf{A}
XX^{-1}
                   Endege WO,
                                Monahan JE;
) T
     Coldegel R.
     WEI; 2001-662795/76.
1....
1....
     Novel isolated nucleic acid molecule associated with cancerous state of
20
     prostate cells and correlating with presence of prostate cancer, useful
FT
     for detecting presence of prostate cancer, stage of prostate cancer -
200
\mathbb{K}^{\mathbb{K}}
     Claim 1; Page 4088; 11750pp; English.
1/3
XX
     The invention relates to an isolated nucleic acid molecule (I) comprising
CC
     a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
\mathbb{C}\mathbb{C}
     specification or its complement. (I) is useful for:
CC
      (a) assessing whether a patient is afflicted with prostate cancer;
CC
      (b) monitoring the progression of prostate cancer in a patient;
CC
      (c) assessing the efficacy of a test compound to inhibit prostate
CC
     cancer in a patient;
CC
      (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC
CC
      in a patient;
      (e) selecting a composition for inhibiting prostate cancer in a patient;
CC
      (f) assessing the prostate cell carcinogenic potential of a compound;
CC
      (g) determining whether prostate cancer has metastasized in a patient;
CC
      (h) assessing the aggressiveness or indolence of prostate cancer in a
CC
CC
      patient;
      (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
CC
XX
      Sequence 3849 BP; 1142 A; 745 C; 858 G; 1081 T; 23 other;
30
                                    Score 582.6; DB 23; Length 3849;
                            57.5%;
  Query Match
                                    Pred. No. 1e-151;
  Best Local Similarity
                            74.7%;
                                    0; Mismatches 244; Indels
                                                                      9; Caps
                                                                                  1;
  Matches 748; Conservative
```

Qy Db		GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAACACAATTTGGC	
Qy	78	CATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTC	128
Db Qy		TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	
Db			
ŌΆ		AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT	
Db		AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT	
Qy Db		AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA	
Q <u>v</u>		CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC	
Db Cy		TOATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG	•
) <u>05</u> -		GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA	
Çy Db	•	TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC	
Qy -		CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA	
Qy		ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT	
Db		ACAGCATTATGATAGATTTTCAGTGAATATGAGAAGTTACTTCATTCA	
Øχ		TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC	
Db Ωy		CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACC'ICCTTCG	
Db			
QУ		GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAG	
Db	1217	AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTTGTAGCCAA	1276

*

```
849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
QΫ
             1277 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 1336
Db
         909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
ΩУ
             1337 GTTCCTCAGCAAGTTTCAGAACGACAGGACGAGGAGGAGGAGGAGAGAC 1396
Db
         969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
QУ
             1397 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1437
כים
RESULT 11
ABV28822
     ABV28822 standard; cDNA; 3849 BP.
ID
ХX
    ABV28822;
AE.
XX
     16-SEP-2002 (first entry)
DT
XX -
     Human prostate expression marker cDNA 28813.
DE
XX
     Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
SVI
                                                                2.0
     pharmacogenomic marker; gene; ss.
W
             7 4.5
ZX
     Homo sapiens.
្ន
XX
     WO200160860-A2.
\mathbb{P}N
\mathbf{X}^{\square}
7.
     23-AUG-2001.
                                                                SX
     20-FEB-2001; 2001WO-US05171.
253
KX.
     17-FEB-2000; 2000US-183319P.
32
     16-MAR-2000; 2000US-189862P.
\mathcal{D}\mathcal{D}
     25-MAY-2000; 2000US-207454P.
ĽR
     09-JUN 2000; 2000US-211314P.
PR
     13-JUL-2000; 2000US-219007P.
υR
     13-DEC-2000: 2000US-255281P.
PR
XX
      (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
ХX
     Schlegel R, Endege WO,
                             Monahan JE;
PI.
XX
     WPI; 2001-662795/76.
 DR
 XX
     Novel isolated nucleic acid molecule associated with cancerous state of
 T\mathbb{C}
     prostate cells and correlating with presence of prostate cancer, useful
 PT
     for detecting presence of prostate cancer, stage of prostate cancer -
 PT
 XX
     Claim 1; Page 6066-6067; 11750pp; English.
 PS
 XX
     The invention relates to an isolated nucleic acid molecule (I) comprising
 CC
      a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC
      specification or its complement. (I) is useful for:
 CC
      (a) assessing whether a patient is afflicted with prostate cancer;
 CC
```

```
(b) monitoring the progression of prostate cancer in a patient;
CC
    (c) assessing the efficacy of a test compound to inhibit prostate
CC
    cancer in a patient;
CC
    (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC
    in a patient;
CC
    (e) selecting a composition for inhibiting prostate cancer in a patient;
CC
    (f) assessing the prostate cell carcinogenic potential of a compound;
CC
    (g) determining whether prostate cancer has metastasized in a patient;
CC
    (h) assessing the aggressiveness or indolence of prostate cancer in a
CC
    patient;
CC
    (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
CC
XX
    Sequence 3849 BP; 1142 A; 745 C; 858 G; 1081 T; 23 other;
SQ
                           Score 582.6; DB 23;
                                             Length 3849;
                     57.5%;
 Query Match
                           Pred. No. 1e-151;
                     74.78;
 Best Local Similarity
                                                        Gaps
                                                               1;
                             Mismatches 244;
                                             Indels
 Matches 748; Conservative
                           0;
         18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
OV
           437 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 496
ďŒ
                                  AAAAAGACAGACAAGGCTTCAGAAGAAG'IGTC 128
         78 CATTTTGGAAAAGCAAGAC
ΰŅ
                                       TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 556
Эb
        Ш
        CAAAAATCTGGTTGCCATGAAAGAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 616
DW
        AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
QY
                                                      1111 11 11
            AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 676.
- in
        AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308.
OY
                  677 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 736
Db
        368 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC
QУ
                737 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 796
Db~
        369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
Q7
                  797 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 856
Db
        429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
QУ
                              857 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 916
Db
         439 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548-
 QΥ
                    917 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 976
 Db
         549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
 Q.Y.
             1111 | 1 | 1 | 1
         977 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA 1036
 Db
```

```
609 ACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
                               11 11111111 11 11111 11 11 11111111
QУ
               1037 ACAGCATTATGATAGATTTTCAGTGAATATGAGAAGTTACTTCAGTAAAATTATGT 1096
Db
        669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
            QУ
       Db
        729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
            Qу
       1157 PATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 1216
Db
        789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAG 848
             QV
       1217 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTTGTAGCCAA 1276
Db
        849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
            Ωy
        1277 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 1336
ďŒ
        909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
            QУ
        1337 GTTCCTCAGCAAGTTTCAGAACGACAGGACGAGAGGAGTGAGCAGTTTAACGACGAGAAGAC 1396
dG
         D65 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
            C_{-}
        1397 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1437
 \mathbf{D}
 RESULT 113
 AAF30688
     AAF30688 standard; cDNA; 1053 BP.
 ID.
 X\Sigma
     AME30688; . .
 AC
 XY
     11-JUN-2001 (first entry)
 DT.
     Human acute neuronal induced calcium binding protein ANIC-BP-1B cDNA.
 200
 125
     Acute neuronal induced calcium binding protein; ANIC-BP-1B;
 XX
     spice variant; human; stroke; head trauma; Parkinson's disease;
 KM
     Alzheimer's disease; multiple sclerosis; spinal cord injury;
 XII
     cerebroprotective; antiparkinsonian; nootropic; neuroprotective;
 KM
 KW
     therapy; diagnosis; vaccine; ss.
 WX
 X
     Homo sapiens.
 OS
 XX
                   Location/Qualifiers
      Key-
 FH
                   1..1053
      CDS
  FT
                   /*.tag= a
  FT
                   /product= "Human ANIC-BP-1B"
  FT
  XX
      WO200125423-A1.
  PΝ
  XX
      12-APR-2001.
  рD
  XX
```

```
28-SEP-2000; 2000WO-EP09475.
PF
XX
                  99EP-0119113.
    04-OCT-1999;
PR
XX
     (MERE ) MERCK PATENT GMBH.
PA
XX
    Duecker K, Den Daas I;
PΙ
XX
    WPI; 2001-266306/27.
DR
     P-PSDB; AAB20387.
DR
XX
    Novel human acute neuronal induced calcium-binding protein like protein
PT
     splice variant, useful for treating stroke, acute head trauma,
PT
     Parkinson's disease, Alzheimer's disease multiple sclerosis, spinal
קים
     cord injury -
\mathbf{p}\mathbf{r}
XX
     Claim 4; Page 43-44; 49pp; English.
PS
XX
     The present sequence is that of cDNA encoding a novel human acute
CC
     neuronal induced calcium binding protein-like protein splice
CC.
     variant, ANIC-NP-1B (see AAB20387). The protein shows homology to
CC
     other members of the calcium binding protein family, including
CC
     ANIC-BP, a protein discovered by mRNA differential display that is
CC
     upregulated in a rat model of head trauma. ANIC-BP and ANIC-BP-1B
CC
     differ in their C-terminal portions. The variant protein could
CC
     Merve as a novel drug target. The invention provides ANIC-BP-1B
CX^{*}
     polynucleotides and polypeptides, expression vectors, host cells
GC.
     and antibodies, as well as methods for producing the protein and
10
     for treating or preventing disorders associated with expression of i \varepsilon
CC
     the protein by inhibiting or activating the action of ANIC-BP-1B. (1)
₽Ç
     seases that may be treated include stroke and acute head trauma, C. .
CC
     workinson's disease, Alzheimer's disease, multiple sclerosis and
CC.
     spinal cord injury. The polynucleotides and polypeptides can also ex
100
     be used in diagnostic assays and in vaccines, and to identify
Ċ
     agonists and antagonists useful for treating conditions associated
ÒĠ
     with ANIC-BP-1B imbalance.
ďC
KX
     Sequence 1053 BP; 357 A; 211 C; 214 G; 271 T; 0 other;
80
                                Score 541.6; DB 22; Length 1053;
                         53.4%;
  Query Match
                         74.1%; Pred. No. 1.5e-140;
  Best Local Similarity
                                0; Mismatches 239; Indels
                                                             11; Gaps
  Matches 716; Conservative
           18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
 QУ
              12 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGCATGGC 71
 ΰ'n
                                        AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
           78 CATTTTGGAAAAGCAAGAC--
 Ωy
                                        11 11111111111
           72 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 131
 Do
          QV
                          | | |
          132 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 191
 Db
          189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
 ØУ
```

	192 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 251
Db	249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308:
Qy -	249 AGCTGACCTGCAGCTGATAGACTTTGAGGGGAAAAAATTTTCAACAA 311 252 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 311
Db	
Qy	309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
Dp	
Qy	369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
Db	
QУ	429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
ďQ	432 AATAATGTTAAGAGAATGCTTCAGATGCTTCAGATGC 548
QУ	489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
DÞ	CTAACCAGACATAAAGTGTTGGTAGCAGACTTAGA 608
Ο.Υ. ,	549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATTATTTTTGA 611
Dh	A CA CTATICA A STATICA CA A A PTICOTT CAGTOTGAGAATTATGT 668)
©y	ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGGTTAATTACGACAAAATTACGACAAGACTATGAGAAAATTACGT 674:
Эb	569 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728.
Q y	69 TACTAAGAGACAGTCTTTAAAGCTGCTTGGTGAACTACTAGATAGA
7424	TEACOGARCECCAGARCECTGARACTCATGATGAACCTCCTTCG 788
QΫ	
DЪ	732 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGTTO
СХ	789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAG 848
Dp	792 AGACAAAGTCGCAACATCCAGTTTGAGGGGTTTGTG
.Q <u>V</u>	849 TCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
Гb	
q_{jj}	909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAG 966
Гb	912 GTTCCTCAGCAAGTTTCAGAACGACAGGACGGATTGTATGAGGACGGATTGTATGA
QУ	967 AACTAC 972
Db	372 GAATIC 977

```
AAS89557 standard; cDNA; 1162 BP.
ID
XX
     AAS89557;
AC
XX
     13-FEB-2002
                  (first entry)
DT
XX
     DNA encoding novel human diagnostic protein #25361.
DΕ
ХX
     Human; chromosome mapping; gene mapping; gene therapy; forensic;
ΚW
     food supplement; medical imaging; diagnostic; genetic disorder; ss.
KW
XX
OS
     Homo sapiens.
XX
     WO200175067-A2.
ΡN
XX.
PD
     11-0CT-2001.
XX
     30-MAR-2001; 2001WO-US08631.
PF
XX
     31-MAR-2000; 2000US-0540217.
PR
     23-AUG-2000; 2000US-0649167.
₽R
XX
     (HYSE-) HYSEQ INC.
\mathbb{P}\mathbf{A}
X.X
     Drmanac RT, Liu C,
                           Tang YT;
PT
XX
                                                                     : F
     WPI; 2001-639362/73.
\partial \mathbb{R}
.....
     P-PSDB; ABG25370.
XX
     New isolated polynucleotide and encoded polypeptides, useful in
PT
     diagnostics, forensics, gene mapping, identification of mutations
22
     responsible for genetic disorders or other traits and to assess
Pi
     biodiversity
MX.
     Claim 1; SEQ ID No 25361; 103pp; English.
PS.
XX
     The invention relates to isolated polynucleotide (I) and
CC
     polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC
     polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC
     and gene mapping, and in recombinant production of (II). The
CC
     polynucleotides are also used in diagnostics as expressed sequence tags
CC
     for identifying expressed genes. (I) is useful in gene therapy techniques
CC
     to restore normal activity of (II) or to treat disease states involving
CC
      (II). (II) is useful for generating antibodies against it, detecting or
CC
     quantitating a polypeptide in tissue, as molecular weight markers and as
CC
     a food supplement. (II) and its binding partners are useful in medical
CC
      imaging of sites expressing (II). (I) and (II) are useful for treating
CC
     disorders involving aberrant protein expression or biological activity.
CC
     The polypeptide and polynucleotide sequences have applications in
CC
     diagnostics, forensics, gene mapping, identification of mutations
CC
      responsible for genetic disorders or other traits to assess biodiversity
CC
      and to produce other types of data and products dependent on DNA and
CC
      amino acid sequences. AAS64197-AAS94564 represent novel human
cc.
      diagnostic coding sequences of the invention.
CC
     Note: The sequence data for this patent did not appear in the printed
CC
      specification, but was obtained in electronic format directly from WIPO
```

at ftp.wipo.int/pub/published_pct_sequences.

CC

CC

Sequence 1162 BP; 383 A; 241 C; 258 G; 280 T; 0 other;

	Query Ma Best Loc Matches	atch— cal Similarity 743; Conserva	53.2%; Score 73.9%; Pred. tive 0; Mis	No. 5.6e-14	10;	14;	Gaps	4;
QΣ		18 GTTTAGTAAAT		11111				
5} 5 <u>7</u>		78 CATTTTGGAAA	1111111				111 11	
ည်း [[d		TAAATCACTGC	11 11111	1111111 111		1111111		
Ω; D:		189 AGAAGCAGTGG	1111 31 11 111			1 1 1		
् ्र	•	249 AGCTGACCTGC				1 11 11		
41		CATCTTGAGAA		1 1 1 1 1 1 1	i {		1 11	
Ç	- (1) - (2) - (3)	TG8 CTCA-TATCO		i!	1 1 1 1		1 1 1	
Ç	in . Y	TGTGGGATTA' 563 TGTGGAATAA'	1					
)y Db	484 TCTAATCAA'I' .	1 1111111				111111	
)y Db	544 GATGCCTTTG 693 GATGCATTTS	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		111111	1 11		
	op Så	604 TTAGAACAAA 743 TTGGAACAGC				1 1 1 1 1 1		
	Op SÀ	554 TATGTTACTA 803 TATGTGACAA	1 1111111	1111 11 11		1 11 1		
	Şγ	724 TTTGCCATCA	TGACAAAGTATATO	AGCAAGCCGGA	GAACCTGAAAC	TCATGAT	GAACCTC	783

```
863 TTCACAATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTG 922
Гb
         784 CTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTG 843
Qy
            923 CTGCGAGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTTGTA 982
Db
         844 GCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTC 903
QУ
             983 GCCAATCCTAACAAGACGCCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTC 1042
         904 ATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAG 963
QУ
             11111
        1043 ATAGAGTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAG 1102
Db
         964 AAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
Çv
             1103 AAGACCTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1148
Db
RESULT 14
AAX39818/c
    AAX39818 standard; DNA; 833 BP.
ID:
XX
AC
     AAX39818;
XX
     ევ-ქლ-1999 (first entry) -
DT
~~ <
     Castric cancer associated gene.
1047:
1113.
     Cancer associated antigen; diagnosis; research; treatment; human;
KW.
     breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW
     préstate cancer; ss.
Y_{ij}^{*}
XX
ČÉ.
     Memo sapiens.
XX
     9:09904265-A2.
ЪИ
XX
     20-JAN-1999.
ОŒ
\Sigma X
                   98WO-US14679.
     15-JUL-1998;
 PБ
XX.
                   98US-0102322.
 p_{R}
     22-JUN-1998;
                   97US-0896164.
 ĖΩ
     17-3UL-1997;
                   97US-0061599.
ÐÑ,
     10-OCT-1997;
     10-OCT-1997;
                   9708-0061765.
 PR
                   97US-0948705.
     10-OCT-1997;
 שט
                   97GB-0021697.
     11-OCT-1997;
 PF.
 XX
      (LUDW-) LUDWIG INST CANCER RES.
 РA
 XX
     Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
 \mathbf{p}\mathbf{I}
     Pfrewidschuh M, Sahin U, Scanlan MJ, Stockert E;
 PI
 PT
     Tureci O;
 XX.
     WPI; 1999-132448/11.
 DŘ
 XX
     New isolated cancer associated nucleic acids and polypeptides -
 PT
```

isolated using sera from cancer patients, used to develop products PT for the diagnosis, monitoring or treatment of cancers PTXXClaim 67; Page 559; 787pp; English. PS The invention relates to a method for diagnosing a disorder characterised XXby expression of a human cancer associated antigen precursor coded for by CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a CC biological sample isolated from a subject with an agent that specifically CC CC binds to the NAM, an expression product or a fragment of an expression product complexed with an HLA molecule; and (b) determining the CC interaction between the agent and the NAM or the expression product as a CC CC determination of the disorder. The products and methods can be used in CC the diagnosis, monitoring, research, or treatment of conditions CCcharacterised by the expression of various cancer associated antigens. CC The invention provides nucleic acid sequences and encoded polypeptides which are cancer associated antigen precursors expressed in human breast CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and CC CC lung cancer. CC XXSequence 833 BP; 253 A; 171 C; 172 G; 227 T; 10 other; SO Length 833; Score 520.2; DB 20; 51.3%; Query Match Pred. No. 1.2e-134; 98.18; Best Local Similarity Gaps Indels Mismatches 7: Matches 566; Conservative 0; 442 GAATGTATTCGACATGAACCACTTG-CCAAAATCATCCTC-TTTTCTAATCAAT,TCAGAG 499 ADD THEFT, IN THE RECORD REPORTED THE 35 732 GAATNTATTCGACTTGACCCANTTGCCCAAANTCATCCTCTTTTTCTAATCAATTCAGAG 673 TO 500 ATTTCTTTAAGT-ACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558 92 ATTTCTTTAAGTAACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 613 Db 559 TTCAAGGATTTACTAACCAGA-CATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTA 617 QV612 TTCAAGGATTTACTAACCNGACCTTAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTA 553 Db 618 CGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAG 677 Q7/ 552 CGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAG 493 DD578 ACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGAC 737 Øу \$92 ACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGAC 433 Db 738 AAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAG 797 Qy432 AAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAG 373 D.b 796 TCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAA 857 Çλ 372 TCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAA 313 Db 858 AACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAG 917

312 AACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAG 253

Qу

Db.

```
918 CAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGAT 977
              QУ
         252 CAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGAT 193
Db
          978 TAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Q7
              192 TAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 156
Dib
RESULT 15
AAS88031
     AAS88031 standard; cDNA; 2492 BP.
ID
XX
     AAS88031;
AC
XX
     13-FEB-2002 (first entry)
DT
     DNA encoding novel human diagnostic protein #23835.
XX.
DE
     Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX
     food supplement; medical imaging; diagnostic; genetic disorder; ss.
 КW
 KW
 XX
     Romo sapiens.
 OS
 XX
      Wall 00175067-A2.
 \mathbb{P}N
 XY
      2:-OCT-2001.
 DD
 X.
                                                                   13
      ....-NAR-2001; 2001WO-US08631.
 85
 #1-MAR-2000; £000US-0540217.
 73-AUG-2000: 2000US-0649167.
 T_{ij}
 \Sigma X
      (HYSE-) HYSEQ INC.
 PΑ
 XX
                           Tang YT;
      Drmanac RT, Liu C,
 PΙ
 XX
      WPI; 2001-639362/73.
 T/Q
      P-PSDB; ABG23844.
 DR
      New isolated polynucleotide and encoded polypeptides, useful in
  ХΧ
      diagnostics, forensics, gene mapping, identification of mutations
  ^{2}\mathrm{T}
      responsible for genetic disorders or other traits and to assess
  φ'n
  PT
      biodiversity
  PT
  XX
       Claim 1; SEQ ID No 23835; 103pp; English.
  PS
       The invention relates to isolated polynucleotide (I) and
  XX
       polypeptide (II) sequences. (I) is useful as hybridisation probes,
  CC
       polymerase chain reaction (PCR) primers, oligomers, and for chromosome
  CC
       and gene mapping, and in recombinant production of (II). The
  CC.
       polynucleotides are also used in diagnostics as expressed sequence tags
  CC
       for identifying expressed genes. (I) is useful in gene therapy techniques
  CC
       to restore normal activity of (II) or to treat disease states involving
  CC
       (II). (II) is useful for generating antibodies against it, detecting or
  CC
       quantitating a polypeptide in tissue, as molecular weight markers and as
  CC
```

CC

```
imaging of sites expressing (II). (I) and (II) are useful for treating
CC
    disorders involving aberrant protein expression or biological activity.
CC
    The polypeptide and polynucleotide sequences have applications in
CC
    diagnostics, forensics, gene mapping, identification of mutations
CC
    responsible for genetic disorders or other traits to assess biodiversity
CC
    and to produce other types of data and products dependent on DNA and
CC
    amino acid sequences. AAS64197-AAS94564 represent novel human
CC
    diagnostic coding sequences of the invention.
CC
    Note: The sequence data for this patent did not appear in the printed
CC
    specification, but was obtained in electronic format directly from WIPO
CC
    at ftp.wipo.int/pub/published_pct_sequences.
CC
XX
    Sequence 2492 BP; 751 A; 477 C; 546 G; 718 T; 0 other;
SQ
                          Score 496; DB 23;
                                            Length 2492;
                     48.9%;
 Query Match
                            Pred. No. 1.1e-127;
 Best Local Similarity
                     73.0%;
                                                               7:
                           0; Mismatches 255;
                                             Indels
                                                        Gaps.
 Matches 737; Conservative
         18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAAGACAATTTGGC 77
QУ
                            1111 1 11 11 1111
        143 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 202
Db
                                  -AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
         78 CATTTTGGAAAAGCAAGAC--
QY
                                   203 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 262
Db.
        ÚУ
                     CAAAAATCTGGTTGCCATGAAAGAATTCTGTATGGCACAAATGAAAAAGATCGTCAGAC 3226
23
        139 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
ŰΥ
            323 AGAAGCAGCAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTATCACCCTGGT 382
Db
        249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 30.8
QΫ
                 383 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 442
Db
        309 CATCTTGAGAAGACAGATAGGCA-CTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATC 367
Q7
               43 TATTCTCAGAAGACAAATTGGTACCGAGAACTCCTACTGTTGAATACATCTGCACCCAAA 502
Db
         358 CTCA--TATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATT--GCCTTACGT 423
CV
                     1 1
         503 CAGAATATTTTTGTTCATGTTATTGAAAGGGTATGAATCTCCCAGAAATAGCTCTAAATT 562
Ďb
         424 TGTGGGATTATGCTGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTT 483
Qу
                                   563 TGTGGAATAATGTTAAGAGAATGCATCAGACATGAACCACTTGGCAAAATCATTTTGTGG 622
Db
         434 TCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCA 543
 ÖХ
                                          | | | | | |
         623 TCGGAACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCA 682
 Db
         544 GATGCC-TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTT 602
 QУ
```

a food supplement. (II) and its binding partners are useful in medical

CC

Db		683	GATGCCATTTGCCACATTCAAGGGTTTACTTACAAGACATAAATTGCTCAGTGCACAATT	, 1~
Qy		603	CTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAA	662
Db		743		802 [.]
Qy		663	TTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAA	722
Db		803		862
Qy		723	CTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	780
Db .				
Qy			THE STANDARD CHECCAROLTCCAGTTTGAAGCCTTTCATGTTTTTAAGG-TGTT	
Db		923	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTIIIIIIIIII	982
Qy.			TOTAL COLOR CTCA CA A A CACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA	
Db		983	TGTGGCCAGTCCTCACAGATCTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOTOT	1042
ÇŲ.			THE STEEL CONCERTS OF A DAMES AND A STREET OF A DAMES AND A STREET OF A DAMES AND A DAMES	
95			ACTCATTGAGTTTCTGAGCAGCTTCCAAATTTTTTTTTT	
δλ	· .		CONTACTACTTCATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009	
Ωy Db			CGAGAAGACTACTTOTTOTTOTTOTTOTTOTTOTTOTTOTTOTTOTTOTTO	J
, 1232				*

Earth completed: January 6, 2004, 01:28:41

GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 01:16:12; Search time 94 Seconds

(without alignments)

4761.303 Million cell updates/sec

Title: US-10-088-872-1

Perfect score: 1014

Sequence: 1 atgaaaaaatgcctttgtt.....tgaagaaaacggccccttga 1014

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA:*

1: /cgn2 6/ptodata/1/ina/5A COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2 6/ptodata/1/ina/PCTUS COMB.seq:*

6: /cgn2 6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

			90				•
Res	sult		Query				
	No.	Score	Match	Length	DB	ID	Description
	1	1014	100.0	1421	 4	US-09-620-312D-111	Sequence 111, App
	2	1010.8	99.7	1344	3	US-09-190-965-2	Sequence 2, Appli
	3	1010.8	99.7	1344	4	US-09-470-253-2	Sequence 2, Appli
С	4	51.6	5.1	7218	1	US-08-232-463-14	Sequence 14, Appl
	5	37.8	3.7	1457	4	US-09-214-307A-9	Sequence 9, Appli
С	6	37.4	3.7	4103	4	US-09-620-312D-390	Sequence 390, App
	7	36.4	3.6	4533	3	US-08-726-214-5	Sequence 5, Appli
С	8	35.6	3.5	577	4	US-09-513-057C-20	Sequence 20, Appl
С	9	35.6	3.5	3707	3	US-09-276-531-42	Sequence 42, Appl
С	10	35.6	3.5	5714	4	US-09-620-312D-393	Sequence 393, App
	11	35.2	3.5	12793	4	US-09-004-838-124	Sequence 124, App

	12	35	3.5	775	3	US-08-961-083-89	Sequence	89, Appl
	13	35	3.5	775	4	US-09-536-784-89	Sequence	89, Appl
	14	35	3.5	1696	3	US-08-961-083-217	Sequence	217, App
	15	35	3.5	1696	4	US-09-536-784-217	Sequence	217, App
	16	35	3.5	3763	1	US-07-792-865D-1	Sequence	1, Appli
С	17	35	3.5	13926	4	US-08-961-527-5	Sequence	5, Appli
С	18	34.8	3.4	1413	4	US-09-016-434-1382	Sequence	1382, Ap
	19	34.8	3.4	12734	4	US-09-344-456-1	Sequence	1, Appli
	20	34.8	3.4	14078	3	US-09-433-262-1	Sequence	1, Appli
	21	34.8	3.4	14078	4	US-09-702-330-1	Sequence	1, Appli
	22	34.8	3.4	14578	3	US-08-859-694-1	Sequence	1, Appli
С	23	34.6	3.4	64467	4	US-09-803-671B-3	Sequence	3, Appli
	24	34.4	3.4	22481	4	US-08-367-841A-43	Sequence	43, Appl
	25	34.4	3.4	22481	5	PCT-US95-07201-43	Sequence	43, Appl
	26	34.4	3.4	22484	4	US-09-875-223-2	Sequence	
	27	34	3.4	3378	4	US-09-328-352-4107	Sequence	4107, Ap
С	28	34	3.4	43360	4	US-09-453-702B-206	Sequence	
С	29	34	3.4	45325	4	US-09-453-702B-261	Sequence	
	30	33.8	3.3	189	4	US-09-134-001C-1145	Sequence	
С	31	33.8	3.3	4233	3	us-09-056-105-27	Sequence	
	32	33.6	3.3	1276	3.	US-09-177-325-2	Sequence	
	33	33.6	3.3	1276	3	US-09-411-812A-2	Sequence	
	34	33.6	3.3	1276	4	US-09-590-113-2	Sequence	
	35	33.6	3.3	1347	4	US-09-134-001C-1309	Sequence	
	36	33.4	3.3	1410	4	US-09-328-352-124	Sequence	
С	37	33.4	3.3	1508	1	US-08-236-311-5	Sequence	
C	38	33.4	3.3	1508	3	US-08-457-918-5	Sequence	
	39	33.4	3.3	2643	4	US-09-486-072-6	Sequence	
С	40	33.4	3.3	18627 -	4	US-08-961-527-113	Sequence	
	41	33.2	3.3	378	4	US-09-252-991A-10313	Sequence	
С	42	33.2	3.3	813	4	US-09-252-991A-10547	Sequence	
	43	33.2	3.3	831	4	US-09-252-991A-10512	Sequence	
C	44	33.2	3.3	3253	4	US-09-333-214-4	Sequence	
	45	33.2	3.3	8091	4	US-09-230-652-1	Sequence	1, Appli

ALIGNMENTS

RESULT 1

US-09-620-312D-111

- ; Sequence 111, Application US/09620312D
- ; Patent No. 6569662
- ; GENERAL INFORMATION:
- ; APPLICANT: Tang, Y. Tom
- ; APPLICANT: Liu, Chenghua
- ; APPLICANT: Asundi, Vinod
- ; APPLICANT: Zhang, Jie
- ; APPLICANT: Ren, Feiyan
- ; APPLICANT: Chen, Rui-hong
- ; APPLICANT: Zhao, Qing A.
- ; APPLICANT: Wehrman, Tom
- ; APPLICANT: Xue, Aidong J.
- ; APPLICANT: Yang, Yonghong
- ; APPLICANT: Wang, Jian-Rui
- ; APPLICANT: Zhou, Ping
- ; APPLICANT: Ma, Yunqing

```
Wang, Dunrui
 APPLICANT:
         Wang, Zhiwei
 APPLICANT:
          John Tillinghast
  APPLICANT:
         Drmanac, Radoje T.
  APPLICANT:
  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
  TITLE OF INVENTION: Polypeptides
  FILE REFERENCE: 784CIP2B
  CURRENT APPLICATION NUMBER: US/09/620,312D
  CURRENT FILING DATE: 2000-07-19
  PRIOR APPLICATION NUMBER: 09/552,317
  PRIOR FILING DATE: 2000-04-25
  PRIOR APPLICATION NUMBER: 09/488,725
  PRIOR FILING DATE: 2000-01-21
 NUMBER OF SEQ ID NOS: 1105
  SOFTWARE: pt FL genes Version 1.0
 SEQ ID NO 111
  LENGTH: 1421
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (217)..(1230)
US-09-620-312D-111
 Query Match
                 100.0%; Score 1014; DB 4;
                                    Length 1421;
 Best Local Similarity
                 100.0%;
                       Pred. No. 4.8e-292;
 Matches 1014; Conservative
                      0; Mismatches
                                                   0;
                                    Indels
                                              Gaps
        1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Qу
         217 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 276
Db
       Qу
         Db
      Qy
         Db
      181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Qу
         397 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 456
Db
      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
Qу
         Db
      301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Qу
         Db
      517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576
      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
Qу
         577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636
Db
      421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
Qy
```

Db	637	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	696
Qу	481	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	540
Db	697	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	756
Qу	541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	600
Db	757	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	816
Qу	601	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	660
Db	817	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	876
Qу	661	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	720
Db	877	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	936
Qу	721	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	780
Db	937	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	996
Qу	781	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	840
Db	997	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	1056
QУ	841	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	.900
Db .	1057	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	1116
Qу	901	CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC	960
Db	1117	$\tt CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC$	1176
Qу		GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014	
Db	1177	GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1230	

US-09-190-965-2

- ; Sequence 2, Application US/09190965
- ; Patent No. 6071721
- ; GENERAL INFORMATION:
- ; APPLICANT: Tang, Y. Tom
- ; APPLICANT: Guegler, Karl J.
- ; APPLICANT: Corley, Neil C.
- ; APPLICANT: Gorgone, Gina A.
- ; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
- ; FILE REFERENCE: PF-0635 US
- ; CURRENT APPLICATION NUMBER: US/09/190,965
- ; CURRENT FILING DATE: 1998-11-13
- ; NUMBER OF SEQ ID NOS: 5
- ; SOFTWARE: PERL Program
- ; SEQ ID NO 2
- ; LENGTH: 1344

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3734805
US-09-190-965-2

Query Match 99.7%; Score 1010.8; DB 3; Length 1344;
```

	cal S	Similarity 99.8%; Pred. No. 4.2e-291; 2; Conservative 0; Mismatches 2; Indels 0; Gaps	0;
QУ	1	ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC	60
Db	124	ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC	183
QУ	61	CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC	120
Db	184	CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC	243
QУ	121	GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	180
Db	244	GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	303
Qy	181	CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG	240
Db	304	CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGGTG	363
Qу	241	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA	300
Db	364	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA	423
Qу	301	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	360
Db	424	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	483
QŅ	361	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	420
Db	484	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	543
Qy	421	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	480
Db	544	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	603
Qу	481	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	540
Db	604	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	663
Qу	541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	600
Db	664	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	723
Qу	601	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	660
Db	724	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	783
QУ	661	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	720
Db	784	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	843

```
721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
Qy
          844 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903
Db
       781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTT 840
Qу
          904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963
Db
       841 GTGGCCAGTCCTCACAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
QУ
           964 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1023
Dh
       901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
Qy
          1024 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083
Db
       961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
QУ
           1084 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1137
Db
RESULT 3
US-09-470-253-2
; Sequence 2, Application US/09470253
; Patent No. 6365371
GENERAL INFORMATION:
  APPLICANT: Tang, Y. Tom
  APPLICANT: Guegler, Karl J.
  APPLICANT:
           Corley, Neil C.
           Gorgone, Gina A.
  APPLICANT:
  TITLE OF INVENTION: CALCIUM BINDING PROTEIN
  FILE REFERENCE: PF-0635 US
  CURRENT APPLICATION NUMBER: US/09/470,253
  CURRENT FILING DATE: 1999-12-22
  PRIOR APPLICATION NUMBER: 09/190,965
  PRIOR FILING DATE: 1998-11-13
  NUMBER OF SEQ ID NOS: 5
  SOFTWARE: PERL Program
 SEQ ID NO 2
   LENGTH: 1344
   TYPE: DNA
   ORGANISM: Homo sapiens
   FEATURE: -
   OTHER INFORMATION: 3734805
US-09-470-253-2
                    99.7%; Score 1010.8; DB 4; Length 1344;
 Query Match
 Best Local Similarity 99.8%; Pred. No. 4.2e-291;
                                                            0;
 Matches 1012; Conservative
                         0; Mismatches
                                       2;
                                          Indels
                                                  0; Gaps
         1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
QУ
           124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183
Db
        Qy
```

Db	184	CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC	243
Qу	121	GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAAGAA	180
Db	244		303
Qу	181	CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG	240
Db	304	CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG	363
Qу	241	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA	300
Db	364	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA	423
Qу	301	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	360
Db	424	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	483
Qу	361	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	420
Db	484	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	543
Qу	421	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	480
Db	544	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	603
Qу	481	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	540
Db	604	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	663
Qу	541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	600
Db	664	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	723
Qу	601	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	660
Db	724	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	783
Qу	661	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	720
Db	784	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	843
QУ	721	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	780
Db	844	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	903
QУ	781	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	840
Db	904	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	963
QУ	841	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	900
Db	964	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	1023
QУ	901	CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC	960
Db	1024	CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGACGGATGATGAGCAGTTCGCTGAC	1083

```
RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
  GENERAL INFORMATION:
    APPLICANT: DORNER, F.
    APPLICANT: SCHEIFLINGER, F.
    APPLICANT: FALKNER, F. G.
    TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
    NUMBER OF SEQUENCES: 52
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Foley & Lardner
      STREET: 1800 Diagonal Road, Suite 500
;
      CITY: Alexandria
      STATE: VA
      COUNTRY: USA
      ZIP: 22313-0299
    COMPUTER READABLE FORM:
     MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/232,463
     FILING DATE:
      CLASSIFICATION: 435
   PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US/07/935,313
     FILING DATE:
     APPLICATION NUMBER: EP 91 114 300.6
     FILING DATE: 26-AUG-1991
   ATTORNEY/AGENT INFORMATION:
     NAME: BENT, Stephen A.
;
      REGISTRATION NUMBER: 29,768
      REFERENCE/DOCKET NUMBER: 30472/114 IMMU
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (703)836-9300
      TELEFAX: (703)683-4109
      TELEX: 899149
  INFORMATION FOR SEQ ID NO: 14:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 7218 base pairs
;
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
     IMMEDIATE SOURCE:
      CLONE: pTZqpt-F1s
US-08-232-463-14
                         5.1%; Score 51.6; DB 1; Length 7218;
  Query Match
  Best Local Similarity 3.6%; Pred. No. 4.2e-05;
```

```
Matches 12; Conservative 196; Mismatches 130; Indels 0; Gaps
                                            0;
       1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Qy
        Db
      Qу
         Db
     Qу
        Db
     181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Qу
          Db
     241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
Qу
        Db
     301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAG 338
Qγ
          Db
RESULT 5
US-09-214-307A-9
; Sequence 9, Application US/09214307A
: Patent No. 6544516
; GENERAL INFORMATION:
 APPLICANT: NEUTEC PHARMA PLC
  TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF INFECTIONS OF GRAM POSITIVE
  TITLE OF INVENTION: COCCI
  FILE REFERENCE: PM 259204
  CURRENT APPLICATION NUMBER: US/09/214,307A
 CURRENT FILING DATE: 1999-01-04
 PRIOR APPLICATION NUMBER: PCT/GB97/01830
; PRIOR FILING DATE: 1997-07-07
  PRIOR APPLICATION NUMBER: GB9614274.0
  PRIOR FILING DATE: 1996-07-06
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
  LENGTH: 1457
  TYPE: DNA
  ORGANISM: Staphylococcus aureus
US-09-214-307A-9
               3.7%; Score 37.8; DB 4; Length 1457;
 Query Match
 Best Local Similarity 47.0%; Pred. No. 0.23;
 Matches 150; Conservative 0; Mismatches 167; Indels 2; Gaps
      430 ATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAAT 489
Qу
        386 AATATGAGAACTGTAGTTGATCGACCTAGAACACAATATAAAAAAGTCGTCTTTAATAAT 445
Db
```

```
490 CAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCC 549
Qу
              Db
        446 TTATTTTATCAATTTAGTAAGGATGCCAACTTTGAACCTATTGCTTGTAGACCCTATCGT 505
        550 TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAA 609
Qу
              506 CCTCAAACAAAGGGTCTGTTGAATCATTAGCTAAATTTGTTGAACAGCGTTTAAGACCA 565
Db
        610 CAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTT 669
Qy
             566 TACGATTATGAATTTTATGATGCTG--TAGAACTTATTGGGCTAGTAAACGATTTATGTC 623
Db
        670 ACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCC 729
Qу
                      624 ACGAATTGAATCACTTAGAAATTTCACAAGCAACAGAACAACGACCTATCGACGTTTTCA 683
Db
Qу
        730 ATCATGACAAAGTATATCA 748
            Db
        684 ATTATGAAGAAAAAGAACA 702
RESULT 6
US-09-620-312D-390/c
; Sequence 390, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
 APPLICANT: Tang, Y. Tom
  APPLICANT: Liu, Chenghua
  APPLICANT: Asundi, Vinod
  APPLICANT: Zhang, Jie
  APPLICANT: Ren, Feiyan
  APPLICANT: Chen, Rui-hong
  APPLICANT: Zhao, Qing A.
  APPLICANT: Wehrman, Tom
  APPLICANT: Xue, Aidong J.
  APPLICANT: Yang, Yonghong
  APPLICANT: Wang, Jian-Rui
  APPLICANT: Zhou, Ping
  APPLICANT: Ma, Yunging
  APPLICANT: Wang, Dunrui
  APPLICANT: Wang, Zhiwei
  APPLICANT: John Tillinghast
  APPLICANT: Drmanac, Radoje T.
  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
  TITLE OF INVENTION: Polypeptides
  FILE REFERENCE: 784CIP2B
  CURRENT APPLICATION NUMBER: US/09/620,312D
  CURRENT FILING DATE: 2000-07-19
  PRIOR APPLICATION NUMBER: 09/552,317
  PRIOR FILING DATE: 2000-04-25
  PRIOR APPLICATION NUMBER: 09/488,725
  PRIOR FILING DATE: 2000-01-21
  NUMBER OF SEQ ID NOS: 1105
 SOFTWARE: pt FL genes Version 1.0
; SEO ID NO 390
   LENGTH: 4103
```

```
TYPE: DNA
   ORGANISM: Homo sapiens
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (104)..(3493)
US-09-620-312D-390
                       3.7%; Score 37.4; DB 4; Length 4103;
 Query Match
 Best Local Similarity 60.2%; Pred. No. 0.53;
         62; Conservative 0; Mismatches 41; Indels 0; Gaps
                                                                   0;
 Matches
          6 AAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAA 65
Qу
                       1 1
        4091 ACAAATGAGAAAGTTTCATTTACCTCAAAAAAATCCAGGCTATACAAACAGACAACTGAA 4032
         Qу
            4031 AGCCACATAGGAAATTTCCGAAACACAAAAGAAAAAGTCTCAC 3989
Db
RESULT 7
US-08-726-214-5
; Sequence 5, Application US/08726214
; Patent No. 6107076
; GENERAL INFORMATION:
    APPLICANT: Tang, Wei-Jen
    APPLICANT: Gilman, Alfred G.
;
    TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;
    TITLE OF INVENTION: AND USES THEREFOR
;
    NUMBER OF SEQUENCES: 31
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Arnold, White & Durkee
     STREET: P.O. Box 4433
     CITY: Houston
     STATE: Texas
     COUNTRY: United States of America
      ZIP: 77210
;
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/726,214
      FILING DATE: Concurrently Herewith
      CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 60/005,498
      FILING DATE: 04-OCT-1995
    ATTORNEY/AGENT INFORMATION:
      NAME: Highlander, Steven L.
      REGISTRATION NUMBER: 37,642
      REFERENCE/DOCKET NUMBER: UTSD: 450
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (512) 418-3000
      TELEFAX: (512) 474-7577
  INFORMATION FOR SEQ ID NO: 5:
```

```
SEQUENCE CHARACTERISTICS:
      LENGTH: 4533 base pairs
;
      TYPE: nucleic acid
;
     STRANDEDNESS: single
     TOPOLOGY: linear
US-08-726-214-5
 Query Match
                       3.6%; Score 36.4; DB 3; Length 4533;
 Best Local Similarity 56.8%; Pred. No. 1.1;
 Matches 67; Conservative 0; Mismatches 51; Indels
                                                       0; Gaps
                                                                  0;
        718 CACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATG 777
Qу
                    1 11 1
       2644 CTCATCGCCACCATCATGCTGGTGCAGGTCAGCCACATGGTGAAGCTGACACTCATGCTG 2703
Db
        778 AACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGG 835
Qy
                          | |
Db
       2704 CTCGTCACAGGCGCCGTGACTGCCATCAACCTGTATGCCTGGTGTCCTGTCTTTGATG 2761
RESULT 8
US-09-513-057C-20/c
; Sequence 20, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND
PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
  CURRENT APPLICATION NUMBER: US/09/513,057C
  CURRENT FILING DATE: 2000-02-24
  NUMBER OF SEQ ID NOS: 35
  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
  LENGTH: 577
   TYPE: DNA
   ORGANISM: Lycopersicon esculentum
US-09-513-057C-20
                       3.5%; Score 35.6; DB 4; Length 577;
 Query Match
 Best Local Similarity 51.2%; Pred. No. 0.64;
                                         79; Indels 0; Gaps
        83; Conservative 0; Mismatches
                                                                  0;
        457 GAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTG 516
Qу
            223 GACCCAAATACCCAAAACACATCTTTACATAGAAATCAAGAGATTTCTGAAGCACACAG 164
Db
        517 GAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACC 576
Qу
                  163 AAGCAAAAAAGATGTATAATTTCACAAAATTACTATTATATTTTTCTGTGATCATGTAAC 104
Db
        577 AGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
QУ
            103 AGGCCTTGTTGGTAAGCACAATAATATGAAGAAAGAGATTAC 62
Db
```

```
US-09-276-531-42/c
; Sequence 42, Application US/09276531
; Patent No. 6183968
  GENERAL INFORMATION:
    APPLICANT: Bandman, Olga
    APPLICANT: Lal, Preeti
    APPLICANT: Hillman, Jennifer L.
    APPLICANT: Yue, Henry
    APPLICANT: Reddy, Roopa
    APPLICANT: Guegler, Karl J.
     APPLICANT: Baughn, Mariah R.
     TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
     TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL
PROLIFERATION
    NUMBER OF SEQUENCES: 134
;
     CORRESPONDENCE ADDRESS:
;
       ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;
       STREET: 3174 PORTER DRIVE
       CITY: PALO ALTO
       STATE: CALIFORNIA
       COUNTRY: USA
       ZIP: 94304
     COMPUTER READABLE FORM:
       MEDIUM TYPE: Floppy disk
       COMPUTER: IBM PC compatible
;
       OPERATING SYSTEM: PC-DOS/MS-DOS
;
       SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;
     CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/276,531
       FILING DATE: Herewith
      CLASSIFICATION:
     PRIOR APPLICATION DATA:
       APPLICATION NUMBER: 60/079,677
;
       FILING DATE: March 27, 1998
;
       CLASSIFICATION:
;
     ATTORNEY/AGENT INFORMATION:
       NAME: Lynn E. Murry, Ph.D.
       REGISTRATION NUMBER: 42,918
       REFERENCE/DOCKET NUMBER: PA-0008 US
     TELECOMMUNICATION INFORMATION:
       TELEPHONE: (650) 855-0555
       TELEFAX: (650) 845-4166
   INFORMATION FOR SEQ ID NO: 42:
     SEQUENCE CHARACTERISTICS:
       LENGTH: 3707 base pairs
;
       TYPE: nucleic acid
       STRANDEDNESS: single
       TOPOLOGY: linear
     IMMEDIATE SOURCE:
       LIBRARY: CERVNOT01
       CLONE: 936117
US-09-276-531-42
                          3.5%; Score 35.6; DB 3; Length 3707;
  Query Match
  Best Local Similarity 51.9%; Pred. No. 1.7;
                                                                             0;
           80; Conservative 0; Mismatches
                                                  74; Indels
                                                                 0: Gaps
```

```
Qу
        445 TGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTC 504
             Db
        3154 TGCTTTCAAAATGTGGAACAAACTAAAATATAAGGCTTTTCTGATAAACTATAAAAATTT 3095
        505 TTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAG 564
Qу
               Db
        3094 AATCAGCACTTGGATCTAATGACATATCTTTATAATACTTCCTCTGCAGATACATTCACT 3035
         565 GATTTACTAACCAGACATAAAGTGTTGGTAGCAG 598
Qy
              Db
        3034 TAGTTCAAACCTTAACATACAAAGTTAGTCTCAG 3001
RESULT 10
US-09-620-312D-393/c
; Sequence 393, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
  APPLICANT: Tang, Y. Tom
  APPLICANT: Liu, Chenghua
  APPLICANT: Asundi, Vinod
  APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
;
  APPLICANT: Chen, Rui-hong
;
  APPLICANT: Zhao, Qing A.
  APPLICANT: Wehrman, Tom
  APPLICANT: Xue, Aidong J.
  APPLICANT: Yang, Yonghong
  APPLICANT: Wang, Jian-Rui
  APPLICANT: Zhou, Ping
  APPLICANT: Ma, Yunging
  APPLICANT: Wang, Dunrui
  APPLICANT: Wang, Zhiwei
  APPLICANT: John Tillinghast
  APPLICANT: Drmanac, Radoje T.
  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
  TITLE OF INVENTION: Polypeptides
  FILE REFERENCE: 784CIP2B
  CURRENT APPLICATION NUMBER: US/09/620,312D
  CURRENT FILING DATE: 2000-07-19
  PRIOR APPLICATION NUMBER: 09/552,317
  PRIOR FILING DATE: 2000-04-25
  PRIOR APPLICATION NUMBER: 09/488,725
  PRIOR FILING DATE: 2000-01-21
  NUMBER OF SEQ ID NOS: 1105
  SOFTWARE: pt FL genes Version 1.0
; SEQ ID NO 393
   LENGTH: 5714
   TYPE: DNA
   ORGANISM: Homo sapiens
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (272)..(4312)
US-09-620-312D-393
 Query Match
                        3.5%; Score 35.6; DB 4; Length 5714;
 Best Local Similarity 51.9%; Pred. No. 2.2;
```

```
Matches 80; Conservative 0; Mismatches 74; Indels 0; Gaps
                                                                      0:
        445 TGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTC 504
Qу
             5233 TGCTTTCAAAATGTGGAACAAACTAAAATATAAGGCTTTTCTGATAAACTATAAAAATTT 5174
Db
        505 TTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAG 564
Qу
              | | | | | | | | | | |
        5173 AATCAGCACTTGGATCTAATGACATATCTTTGTAATACTTCCTCTGCAGATACATTCACT 5114
Db
        565 GATTTACTAACCAGACATAAAGTGTTGGTAGCAG 598
QУ
             5113 TAGTTCAAACCTTAACATACAAAGTTAGTCTCAG 5080
Db
RESULT 11
US-09-004-838-124
; Sequence 124, Application US/09004838
; Patent No. 6350933
; GENERAL INFORMATION:
    APPLICANT: Michelmore, Richard W.
    APPLICANT: Shen, Kathy
    APPLICANT: Meyers, Blake
    TITLE OF INVENTION: Procedures and Materials for
    TITLE OF INVENTION: Conferring Pest Resistance in Plants
    NUMBER OF SEQUENCES: 140
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Townsend and Townsend and Crew LLP
;
      STREET: Two Embarcadero Center, Eighth Floor
;
      CITY: San Francisco
      STATE: California
     COUNTRY: USA
      ZIP: 94111-3834
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
;
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/09/004,838
      FILING DATE: 09-JAN-1998
      CLASSIFICATION: 800
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/781,734
      FILING DATE: 10-JAN-1997
    ATTORNEY/AGENT INFORMATION:
      NAME: Einhorn, Gregory P.
      REGISTRATION NUMBER: 38,440
      REFERENCE/DOCKET NUMBER: 023070-078810US
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (415) 576-0200
      TELEFAX: (415) 576-0300
   INFORMATION FOR SEQ ID NO: 124:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 12793 base pairs
;
      TYPE: nucleic acid
;
      STRANDEDNESS: single
```

```
TOPOLOGY: linear
;
    MOLECULE TYPE: DNA
    FEATURE:
      NAME/KEY: -
      LOCATION: 1..12793
      OTHER INFORMATION: /note= "RG2S"
US-09-004-838-124
 Query Match
                       3.5%; Score 35.2; DB 4; Length 12793;
 Best Local Similarity 47.6%; Pred. No. 4.4;
 Matches 101; Conservative 10; Mismatches
                                            98; Indels
        438 GAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAG 497
Qy
            Db
        5998 GAGARAGWAWGRRRGAKAKARMCSMSYTTGGGATGTGATACTTCTTTTAGGAAAATGGAG 6057
        498 AGATTTCTTTAAGTACGTGGAGTTGTCA---ACATTTGATATTGCTTCAGATGCCTTTGC 554
Qy
              Db
       6058 TTATATCTTTGATATTGTATTTTTTAATGTAATTTATATATTTTAATCATTTTAGTTTAT 6117
Qу
        555 TACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAA 614
             6118 AAGTTTTATTTATTTTGATATGAAAAAAAAAGTCTTTTATACATTGGATTTAACATAAAA 6177
Db
        615 TTACGACACTATTTTTGAAGACTATGAGAAAT 646
Qу
             Db
       6178 ATCCAACAATATTAATCAAAAAGACCAMACAT 6209
RESULT 12
US-08-961-083-89
; Sequence 89, Application US/08961083
; Patent No. 6159469
  GENERAL INFORMATION:
    APPLICANT: Choi et. al.
    TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
    NUMBER OF SEQUENCES: 452
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Human Genome Sciences, Inc.
      STREET: 9410 Key West Avenue
      CITY: Rockville
      STATE: Maryland
      COUNTRY: USA
      ZIP: 20850
;
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
      COMPUTER: HP Vectra 486/33
      OPERATING SYSTEM: MSDOS version 6.2
      SOFTWARE: ASCII Text
    CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/961,083
      FILING DATE:
      CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
    APPLICATION NUMBER:
      FILING DATE:
    ATTORNEY/AGENT INFORMATION:
```

```
NAME: Brookes, A. Anders
     REGISTRATION NUMBER: 36,373
     REFERENCE/DOCKET NUMBER: PB340P2
    TELECOMMUNICATION INFORMATION:
     TELEPHONE: (301) 309-8504
     TELEFAX: (301) 309-8512
  INFORMATION FOR SEQ ID NO: 89:
    SEQUENCE CHARACTERISTICS:
     LENGTH: 775 base pairs
     TYPE: nucleic acid
     STRANDEDNESS: double
     TOPOLOGY: linear
US-08-961-083-89
                       3.5%; Score 35; DB 3; Length 775;
 Query Match
 Best Local Similarity 46.9%; Pred. No. 1.1;
 Matches 143; Conservative 0; Mismatches 160; Indels
                                                        2; Gaps
         22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATCCTGAAAGACAATTTGGCCATT 81
Qу
            263 AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 322
Db
         82 TTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
QУ
                                    1 11
        323 GCGCCAAGAGAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 382
Db
        142 GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCCAACAGAAGCAGTGGCT 201
Qу
            383 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 442
Db
        202 CAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
Qу
                443 GAAGAAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAG 500
Db
        262 CTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
QУ
             501 TTGAAACGCCTGCTGTAGAAAAACAAACAAACAAACCAACAGAGGAACCAAAAGTTGAACAAG 560
Db
        322 CAGAT 326
Qу
            11111
        561 CAGGT 565
Db
RESULT 13
US-09-536-784-89
; Sequence 89, Application US/09536784
; Patent No. 6573082
   GENERAL INFORMATION:
       APPLICANT: Choi et. al.
       TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
       NUMBER OF SEQUENCES: 452
       CORRESPONDENCE ADDRESS:
            ADDRESSEE: Human Genome Sciences, Inc.
            STREET: 9410 Key West Avenue
            CITY: Rockville
            STATE: Maryland
            COUNTRY: USA
```

```
ZIP: 20850
       COMPUTER READABLE FORM:
           MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
           COMPUTER: HP Vectra 486/33
           OPERATING SYSTEM: MSDOS version 6.2
           SOFTWARE: ASCII Text
       CURRENT APPLICATION DATA:
           APPLICATION NUMBER: US/09/536,784
            FILING DATE: 30-Oct-1997
           CLASSIFICATION: <Unknown>
       PRIOR APPLICATION DATA:
           APPLICATION NUMBER: 08/961,083
            FILING DATE: OCT-30-1997
       ATTORNEY/AGENT INFORMATION:
           NAME: Michelle S. Marks
            REGISTRATION NUMBER: 41,971
            REFERENCE/DOCKET NUMBER: PB340P3
       TELECOMMUNICATION INFORMATION:
           TELEPHONE: (301) 309-8504
           TELEFAX: (301) 309-8512
   INFORMATION FOR SEQ ID NO: 89:
     SEQUENCE CHARACTERISTICS:
            LENGTH: 775 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: double
            TOPOLOGY: linear
       SEQUENCE DESCRIPTION: SEQ ID NO: 89:
US-09-536-784-89
                      3.5%; Score 35; DB 4; Length 775;
 Query Match
 Best Local Similarity 46.9%; Pred. No. 1.1;
 Matches 143; Conservative 0; Mismatches 160; Indels
                                                       2; Gaps
         22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT 81
Qу
            263 AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 322
Db
         82 TTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
Qy
             323 GCGCCAAGAGAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 382
Db
        142 GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTGGCT 201
Qу
            383 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 442
Db
        202 CAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
Qу
             443 GAAGAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAG 500
Db
        262 CTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
Qу
             501 TTGAAACGCCTGCTGTAGAAAAACAAACAGAACCAACAGAGGAACCAAAAGTTGAACAAG 560
Db
        322 CAGAT 326
Qу
            11111
Db
        561 CAGGT 565
```

```
RESULT 14
US-08-961-083-217
; Sequence 217, Application US/08961083
; Patent No. 6159469
  GENERAL INFORMATION:
    APPLICANT: Choi et. al.
    TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
    NUMBER OF SEQUENCES: 452
;
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Human Genome Sciences, Inc.
      STREET: 9410 Key West Avenue
      CITY: Rockville
      STATE: Maryland
      COUNTRY: USA
      ZIP: 20850
    COMPUTER READABLE FORM:
;
      MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;
      COMPUTER: HP Vectra 486/33
;
      OPERATING SYSTEM: MSDOS version 6.2
;
      SOFTWARE: ASCII Text
    CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/961,083
      FILING DATE:
      CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:
      FILING DATE:
    ATTORNEY/AGENT INFORMATION:
      NAME: Brookes, A. Anders
      REGISTRATION NUMBER: 36,373
      REFERENCE/DOCKET NUMBER: PB340P2
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (301) 309-8504
      TELEFAX: (301) 309-8512
  INFORMATION FOR SEQ ID NO: 217:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 1696 base pairs
;
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
US-08-961-083-217
                        3.5%; Score 35; DB 3; Length 1696;
 Ouery Match
  Best Local Similarity 46.9%; Pred. No. 1.7;
 Matches 143; Conservative 0; Mismatches 160; Indels
                                                          2; Gaps
                                                                  1;
          22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT 81
Qу
                                         275 AGTCAACCATCAGACAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 334
Db
         82 TTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
Qу
                                    335 GCGCCAAGAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 394
Db
        142 GCAATGAAAGAATTCTGTGTGGTACAAACGAGAAGAACCCCCAACAGAAGCAGTGGCT 201
Qу
```

```
Db
         395 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 454
         202 CAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
Qу
                  111
Db
         455 GAAGAAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAG 512
         262 CTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
Qу
              513 TTGAAACGCCTGCTGTAGAAAAACAAACAGAACCAACAGAGGAACCAAAAGTTGAACAAG 572
         322 CAGAT 326
Qу
             Db
         573 CAGGT 577
RESULT 15
US-09-536-784-217
; Sequence 217, Application US/09536784
; Patent No. 6573082
   GENERAL INFORMATION:
        APPLICANT: Choi et. al.
        TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
        NUMBER OF SEQUENCES: 452
        CORRESPONDENCE ADDRESS:
             ADDRESSEE: Human Genome Sciences, Inc.
             STREET: 9410 Key West Avenue
             CITY: Rockville
             STATE: Maryland
             COUNTRY: USA
             ZIP: 20850
        COMPUTER READABLE FORM:
             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
             COMPUTER: HP Vectra 486/33
             OPERATING SYSTEM: MSDOS version 6.2
             SOFTWARE: ASCII Text
        CURRENT APPLICATION DATA:
             APPLICATION NUMBER: US/09/536,784
             FILING DATE: 30-Oct-1997
             CLASSIFICATION: <Unknown>
        PRIOR APPLICATION DATA:
             APPLICATION NUMBER: 08/961,083
             FILING DATE: OCT-30-1997
        ATTORNEY/AGENT INFORMATION:
             NAME: Michelle S. Marks
             REGISTRATION NUMBER: 41,971
             REFERENCE/DOCKET NUMBER: PB340P3
        TELECOMMUNICATION INFORMATION:
             TELEPHONE: (301) 309-8504
             TELEFAX: (301) 309-8512
   INFORMATION FOR SEQ ID NO: 217:
        SEQUENCE CHARACTERISTICS:
             LENGTH: 1696 base pairs
             TYPE: nucleic acid
             STRANDEDNESS: double
             TOPOLOGY: linear
        SEQUENCE DESCRIPTION: SEQ ID NO: 217:
US-09-536-784-217
```

Best Lo	cal :	3.5%; Score 35; DB 4; Length 1696; Similarity 46.9%; Pred. No. 1.7; B; Conservative 0; Mismatches 160; Indels 2; Gaps 1	
Macches	14.	o, conservative o, mismatches 100, indexs 2, sups 1	′
Qу	22	AGTAAATCACACAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT 81	
Db	275	AGTCAACCATCAGACAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 334	
Qy	82	TTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141	
Db	335	GCGCCAAGAGAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 394	
Qу	142	GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAGAACCCCCCAACAGAAGCAGTGGCT 201	
Db	395	GAAGAGAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 454	
Qу	202	CAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261	
Db	455	GAAGAAACTGTAGAACCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAG 512	
Qу	262	CTGAŢĄGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321	
Db	513	TTGAAACGCCTGCTGTAGAAAACAAACAGAACCAACAGAGGAACCAAAAGTTGAACAAG 572	
Qy	322	CAGAT 326	
Db	573	CAGGT 577	

Search completed: January 6, 2004, 03:19:48 Job time: 96 secs

GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

January 6, 2004, 02:35:04; Search time 1394 Seconds Run on:

(without alignments)

2517.743 Million cell updates/sec

US-10-088-872-1 Title:

1014 Perfect score:

1 atgaaaaaatgcctttgtt.....tgaagaaaacggccccttga 1014 Secuence:

IDENTITY_NUC Scoring table:

Gapop 10.0 , Gapext 1.0

2263443 seqs, 1730637950 residues Searched:

Total number of hits satisfying chosen parameters: 4526836

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Published Applications NA:* Database .

/cgn2 6/ptcdata/1/pubpna/US07_PUBCOMB.seq:*

/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:* 4:

/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:* 5:

/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:* 7:

/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cqn2 6/ptodata/1/pubpna/US09A PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2 6/ptodata/1/pubpna/US09C PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*

/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Query Result Score Match Length DB

No.

옿

Description

US-10-117-722-111 Sequence 111, App 1 1014 100.0 1421 13 Sequence 111, App US-10-037-270-111 15 1014 100.0 1421 2 Sequence 2, Appli US-10-025-730-2 1344 15 1010.8 99.7 3 US-09-918-995-5343 Sequence 5343, Ap 475 11 398 39.3 4 Sequence 318, App US-09-910-943-318 28.5 690 9 5 238.8 Sequence 5263, Ap US-09-867-701-5263 24.3 435 10 246.4 6 Sequence 5899, Ap 447 10 US-09-867-701-5899 7 24.1 244.8 Sequence 4953, Ap 450 1.0 US-09-867-701-4953 24.1 8 244.8 Sequence 35, Appl 762 9 US-09-910-943-35 20.8 9 210.8 Sequence 1453, Ap 387 10 US-09-954-456-1453 195 19.2 10 С Sequence 481, App 387 10 US-09-880-107-481 195 19.2 11 Sequence 118, App 16.7 722 1.3 US-10-257-826A-118 1.2 169.8 US-10-257-826A-119 Sequence 119, App 700 13 166.6 16.4 13 Sequence 592, App US-09-770-445-592 15.4 861 9 156 14 Sequence 1251, Ap 7.3 262 9 US-09-923-876-1251 15 74.2 Sequence 1251, Ap US-09-923-876-1251 7.3 262 12 74.2 16 Sequence 19069, A US-09-918-995-19069 6.5 336 11 65.6 17 US-10-242-355-323 Sequence 323, App 6.4 187 12 418 65 Sequence 54, Appl US-10-080-254-54 65 5.4 487 13 1.9 Sequence 13369, A 254 10 US-09-878-574-13369 5.3 20 53.6 Sequence 777, App US-09-770-961-777 486 11 5.0 21 50.6 C Sequence 2528, Ap. US-09-923-876-2528 9 . 41.6 4.1 242 22 Sequence 2528, Ap US-09-923-876-2528 23 41.6 4.1 242 12 Sequence 294, App US-10-310-154-294 1295 12 40.8 4.0 24 Sequence 1007, Ap 4.0 113306 US-10-292-798-1007 2.2 25 40.2 Sequence 5787, Ap-431 .US-09-918-995-5787 11 ∴6 39.8 3.9 US-10-311-455-26 Sequence 26, Appl: 6301 13 1.7 50.4 3.8 Sequence 261235,: US-10-027-632-261235 1200 13 -835.2 3.8 Sequence 261235, ... US-10-027-632-261235 1200 14 29 38.2 3,8 Sequence 9, Appli US-10-054-968-9 15 1457 30 37.8 3.7 Sequence 278, App US-09-873-367C-278 7178 13 01 37.8 3.7 Sequence 45, Appl US-10-001-843-45 3.7 1267 14 32 37.5 Sequence 1, Appli 3.7 3673778 1.3 US-10-312-841-1 33 37.6 Sequence 45, Appl US-10-087-464-45 2232 1.5 3.7 34 37.4 Sequence 335, App US-09-876-889-335 4012 9 35 37.4 3.7 C US-10-117-722-390 Sequence 390, App 4103 13 3.6 37.4 3.7 C Sequence 390, App US-10-037-270-3904103 15 37 37.4 3.7 \mathbb{C} Sequence 1760, Ap US-10-311-455-1760 8577 13 .37.4 3.7 38 5 Sequence 261978,. US-10-027-632-261978 869 13 3.7 39 3.7.2 €. Sequence 261978, US-10-027-632-261978 14 37.2 3.7 869 C 40 Sequence 538, App US-10-311-455-538 3.6 5413 13 41 37 $^{\rm C}$ Sequence 944, App US-10-311-455-944 3.6 9367 1.3 42 36.6 Sequence 143, App $1\dot{4}$ US-10-079-623-143 3.6 461 43. 36.4 Sequence 29299, A 2641 12 US-10-369-493-29299 3.5 124 36.4 Sequence 297, App 6071. 13 US-10-311-455-297 3.6 45 36.4

ALIGNMENTS

RESULT 1 US-10-117-722-111

[;] Sequence 111, Application US/10117722

[;] Publication No. US20030219744A1

[;] GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

```
Liu, Chenghua
 APPLICANT:
 APPLICANT:
          Asundi, Vinod
 APPLICANT:
          Zhanq, Jie
          Drmanac, Radoje T.
 APPLICANT:
 TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
 TITLE OF INVENTION: Polypeptides
 FILE REFERENCE: 784CIP2BCIP
 CURRENT APPLICATION NUMBER: US/10/117,722
 CURRENT FILING DATE: 2002-04-04
 PRIOR APPLICATION NUMBER: 09/620,312
 PRIOR FILING DATE: 2000-07-19
 PRIOR APPLICATION NUMBER: 09/552,317
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: 09/488,725
 PRIOR FILING DATE: 2000-01-21
 NUMBER OF SEQ ID NOS: 1104
 SOFTWARE: pt_FL_genes Version 1.0
 SEQ ID NO 111
  LENGTH: 1421
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (217) .. (1230)
0.8-2.0-117-722-111
                                       Length 1421;
                        Score 1014;
                                 DB 1.3;
                  100.0%;
 Cuery Match
                        Pred. No. 1.2e-281;
                  100.9%;
 West Local Similarity
                                                Gaps ...
                         Mismatches
                                    0; Indels
                       0;
 Matches 2014;
           Conservative
        1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
OY
          217 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 27.6
Db
       Qy
          T) n
       CY
          D.b
       131 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
QY
          397 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 456
∂b
       241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
Qy
          457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 516
Db
       301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
Qy
          517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576
Dρ
       361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
Çу
          577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636
Db
```

Qy			CGTTGTGGGATTATGCTGAGAGATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
Db			CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCGAAATTTTTTTT
ДУ			TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
Dp			TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGGT
QY			TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
Db			TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGATGTAGATGTTGGTAGATGTTGGTAGATGTTGGTAGATGTTGGTAGATGTTGGTAGATGTTGGTAGATGTTGGTAGATGTTGGTAGATGTTGGTAGATGTAGATGTTGGTAGATGTTGGTAGATGTTGGTAGATGTTGGTAGATGTTGGTAGATGTTGGTAGATGTAGATGTTGGTAGATGTAGATGTTGGTAGATGTAGATGTTGGTAGATGAT
Öň			TTCTTAGAACAAAA'ITACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
áQ			TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTGTCTG
Qу			AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
ंव			
QУ			AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
DЪ		937	
QY			OTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
Db			
307	* .		GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
			GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATTCAAGGGGAAA
(2) ·			CTCATTGAGTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
, Db,			7 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGGAAGGACGGATGATGAGCAGTTTTTTTT
Qy			1 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
ďФ		1177	7 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1230

US-10-037-270-111

- ; Sequence 111, Application US/10037270
- ; Publication No. US20030104529A1
- ; GENERAL INFORMATION:
- ; APPLICANT: Tang, Y. Tom
- APPLICANT: Liu, Chenghua
- ; APPLICANT: Asundi, Vinod
- ; APPLICANT: Zhang, Jie
- ; APPLICANT: Ren, Feiyan
- APPLICANT: Chen, Rui-hong
- APPLICANT: Zhao, Qing A.
- ; APPLICANT: Wehrman, Tom
- ; APPLICANT: Xue, Aidong J.
- ; APPLICANT: Yang, Yonghong

```
Wang, Jian-Rui
 APPLICANT:
         Zhou, Ping
 APPLICANT:
 APPLICANT:
         Ma, Yunqing
         Wang, Dunrui
 APPLICANT:
         Wang, Zhiwei
 APPLICANT:
         Tillinghast, John
 APPLICANT:
 APPLICANT: Drmanac, Radoje T.
 TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
 TITLE OF INVENTION: Polypeptides
 FILE REFERENCE: 784CIP2B
 CURRENT APPLICATION NUMBER: US/10/037,270
 CURRENT FILING DATE: 2002-01-04
 PRIOR APPLICATION NUMBER: 09/552,317
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: 09/488,725
 PRIOR FILING DATE: 2000-01-21
 NUMBER OF SEQ ID NOS: 1104
 SOFTWARE: pt_FL_genes Version 1.0
 SEC ID NO 111
  LENGTH: 1421
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/LEY: CDS
  LOCATION: (217) .. (1230)
08-10-037-270-111 ·
                                DB 15; Length 1421;
                        Score 1014;
                  100.08;
 ്യലു y Match
                 100.0%; Pred. No. 1.2e-281;
 Rest Local Similarity
                      0; Mismatches
                                      Indels
                                               .Gaps
                                                     0.:
                                   0;
 Moliches 1014; Conservative
        ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
         ATGANAAAAATCCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 276
ث.ل
       QΖ.
         Eb
       QУ
         Db
       131 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAACTCTACAGCAGTGGCCTGCTAGTG 240
Qy
         CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 456
Do
       241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300.
Qу
          457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 516
Db
       301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
ÇУ
          517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576
Db
       361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
Qy
```

Db		577	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	030
QΥ			CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	
Db.		637	CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC	030
QУ			TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	
Db		697	TTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	756
ΩY			TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	
Db		757	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	816
Ğλ			TTCTTAGAACAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	
Db			TTCTTAGAACAAAATTACGACACTATTTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	
Qy			AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTGGGGGGGG	
Dþ			AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	
ΩΣ			AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	1 1
ØĎ			ACTTTGCCATCATGACAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	
Qy	٠.		CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	
			GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	
ĞΣ.				
). 			GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA	
O.			CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC	
Db			CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC	. 11/0
СУ	,		GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAÁGAAAACGGCCCCTTGA 1014	
Dр		1177	7 GAGAAGAACTACTTGATTAÁACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1230	

US-10-025-730-2

- ; Sequence 2, Application US/10025730
- ; Publication No. US20030045466A1
- ; GENERAL INFORMATION:
 - APPLICANT: Tang, Y. Tom
- APPLICANT: Guegler, Karl J.
- APPLICANT: Corley, Neil C.
- APPLICANT: Gorgone, Gina A.
 - TITLE OF INVENTION: CALCIUM BINDING PROTEIN
- ; FILE REFERENCE: PF-0635 US
- ; CURRENT APPLICATION NUMBER: US/10/025,730
- ; CURRENT FILING DATE: 2001-12-18
- PRIOR APPLICATION NUMBER: US/09/190,965

```
PRIOR FILING DATE: 1998-11-13
     NUMBER OF SEQ ID NOS: 5
     SOFTWARE: PERL Program
   SEQ ID NO 2
       LENGTH: 1344
       TYPE: DNA
       ORGANISM: Homo sapiens
       FEATURE: -
       OTHER INFORMATION: 3734805
US-10-025-730-2
                                                                  Score 1010.8; DB 15; Length 1344;
                                                   99.7%;
   Query Match
                                                                  Pred. No. 1e-280;
                                                   99.2%;
   Best Local Similarity
                                                                                                                                                        0:
                                                                                                                                       Gaps
                                                                        Mismatches
                                                                                                    2;
                                                                                                            Indels
   Matches 1012; Conservative
                       1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Ωy
                            124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183
Db
                    QY
                            Do
                    OV
                            Э'n
                    ±81 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
\Omega V
                            194 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGCTG 363
 מת
                    ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
                            ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423 4
 Cb
                    201 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
 QУ
                             424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483
 100
                    361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 QY
                             #84 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543
 Ūβ
                    421 \quad \text{CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC} \quad 480 \cdot 100 
 ⊋у
                             564 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603
 Db
                     481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 Qу
                             504 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663
 D_{ij}
                     541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
 QУ
                             664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 723
  Db
                     601 TTCTTAGAACAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
  Qy
```

```
724 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783
Db
        661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
Qy
            734 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843
Db
        721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
Qу
            844 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903
Db
        781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
Qу
            904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963
Db
        841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
Qv
            964 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1023
Db
        901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
QУ
            1024 CICATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCIGAC 1083
Db
         961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
QΥ
            1994 CAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1137
Db
ARSULT 4
                                                              1 %
US-09-012-995-5343
                                                               1 617
Sequence 343, Application US/09918995
 Pablication No. US20030073623A1
                                                           45.1.4
 TOTMERAL INFORMATION:
  APPLICANT: Hyseq, Inc.
   TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
  TITLE OF INVENTION: FROM VARIOUS CONA LIBRARIES
   FILE REFERENCE: 20411-756
   CURRENT APPLICATION NUMBER: US/09/918,995
   CURRENT FILING DATE: 2001-07-30
   PRIOR APPLICATION NUMBER: US/09/235,076
   PRIOR FILING DATE: 1999-01-20
   NUMBER OF SEQ ID NOS: 38054
   SOFTWARE: FastSEQ for Windows Version 3.0
  SEC ID NO 5343
    LENGTH: 475
    TYPE: DNA
    ORGANISM: Homo sapiens
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (1)...(475)
    OTHER INFORMATION: n = A, T, C or G
US-09-918-995-5343
                       39.3%; Score 398; DB 11; Length 475;
  Query Match
                      100.0%; Pred. No. 3.5e-104;
  Best Local Similarity
                             0; Mismatches
                                            0;
                                               Indels
                                                           Gaps
  Matches 398; Conservative
         617 ACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGA 676
```

QУ

, E

· ij. .

* h.

```
1 ACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGA 60
Dh
        677 GACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGA 736
Qy
           61 GACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGA 120
Db
        737 CAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAA 796
Qу
           121 CAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAA 180
Db
        797 GTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACA 856
Çу
           181 GTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACA 240
Db
        857 AAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGA 916
QУ
           241 AAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGA 300
Db
        917 CCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGA 976
ÖΆ
           301 GCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGA 360
Db
        977 TTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Ωy
           361 TTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 398
25
RESULT 5
US-09-910-043-318
quence 318, Application US/09910943
                                                       Tai.
  ##tent No. US20020081610A1
                                                       15861...
  ENTOURAL INFORMATION:
   APPLICANT: Hemmati-Brivanlou, Ali
  **PLICANT: Altman, Curtis .
  THILE OF INVENTION: Assays and Materials for Embryonic Gene Expression
   FILE REFERENCE: 7529/1G148US1
   CURRENT APPLICATION NUMBER: US/09/910,943
   CURRENT FILING DATE: 2001-07-23
   NUMBER OF SEQ ID NOS: 742
   SOFTWARE: PatentIn version 3.1
 SEC ID NO 318
   LENGTH: 690
   TYPE: DNA
    ORGANISM: Xenopus laevis
    FEATURE:
    NAME/KEY: misc feature
    LOCATION: (1)..(690)
    OTHER INFORMATION: n may be a or g or c or t/u
 US-09-910-943-318
                     28.5%; Score 288.8; DB 9;
                                            Length 690;
  Query Match
  Best Local Similarity 80.5%; Pred. No. 1.4e-72;
                                                              0;
                           0; Mismatches 82;
                                            Indels
                                                        Gaps
 Matches 338; Conservative
         595 GCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAG 654
 QУ
```

```
69 GCAGAATTTCTAGAGCAAAATTACGACAGAATATTTAATGACTATGAAAAGCTTCTTCAC 128
Db
        655 TCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGAC 714
           QУ
        129 TCTGAGAACTATGTGACGAAGAGACAGTCCCTTAAGCTGCTGGGCGAGCTGATCCTGGAC 188
Db
        715 CGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATG 774
           Qу
        189 CGACAACTTTTCCATTATGACTAAATACATAAGCAAGCCTGAAAATCTGAAGCTCATG 248
Db
        775 ATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAG 834
           Qу
        249 ATGAATCTGCTCCGTGATAAGAGCCCAAACATTCAGTTTGAAGCATTCCATGTGTTTAAG 308
Db
        835 GTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAG 894
            Qу
        309 GTGTTTGTAGCAAATCCAAACAAAACACAGCCCATCGTGGATATCCTGTTAAAAAACCAA 368
Db
        .QY
        369 ACCAAGTTAATCGACTTCCTGAGCAGCTTTCAGAAGGATCGAACAGATGACGAACAGTTC 428
Dр
        955 GCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
             Qy
        *30 ACCGACGAGAAGAACTACTTGATCAAACAGATACGAGACTTAAAAAAGCCCACGCCATGA 488
Db
                                                             S
RESULT
                                                            6-
ປຣະນອ-867 -701-5263
 ; Sequence 5263, Application US/09867701
                                                         (). C
 / Regent No. US20020132237A1
                                                         18.0
 PERAL INFORMATION:
  APPLICANT: Aglate, Paul A.
   APPLICANT: Jones, Robert
   APPLICANT: Harlocker, Susan L.
   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
   TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
   FILE REFERENCE: 210121.497
   CURRENT APPLICATION NUMBER: US/09/867,701
   CURRENT FILING DATE: 2001-05-29
   NUMBER OF SEQ ID NOS: 10912
   SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 5263
    LENGTH: 435
    TYPE: DNA
    ORGANISM: Homo sapien
 US-09-867-701-5263
                            Score 246.4; DB 10; Length 435;
                      24.3%;
   Query Match
                            Pred. No. 1.8e-60;
                      77.6%;
   Best Local Similarity
                                              Indels
                            0; Mismatches 86;
   Matches 298; Conservative
         626 TTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTT 685
                   34 1141144 14 11414 1434 1441411 14 14 1414144
 Q٧
          41 TTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGTGACAAAAAGACAGTCAC 100
 ďŒ
         686 TAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATA 745
 Qу
```

```
101 TGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCACAATTATCACAAAATACA 160
Db
        746 TCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACA 805
QУ
           161 TCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAACA 220
Dh
        806 TCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGC 865
\Omega Y
           221 TCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAATCCTAACAAGACGCAGC 280
D'n
        866 CTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCC 925
Qy
           231 CCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGAGTTCCTCAGCAAGTTTC 340
nb
        925 AAAAAGAAAGCACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGA 985
Qv
           341 AGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACCTATTTAGTTAAACAGA 400
        986 TCCCAGACTTGAAGAAAACGGCCC 1009
           401 TCCGGGATTTGAAGAGACCCGCTC 424
Db
RESOLT 7
08-09-887-7-11-5899
/ Sequence [699, Application US/09867701
 Patent No. - US20020132237A1
; GENERAL INFORMATION:
  PLICANT: Aglate, Paul A.
  PPINICANT: Jones, Robert
  WARLICANT: Harlocker, Susan L.
  MINTLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
  WITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
: PIBE REFERENCE: 210121.497
  CURRENT APPLICATION NUMBER: US/09/867,701
  CURRENT FILING DATE: 2001-05-29
  NUMBER OF SEQ ID NOS: 10912
  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5899
   LENGTH: 447
   YYVH: ONA
   ORGANISM: Homo sapien
US-09-867-701-5899
                     24.1%; Score 244.8; DB 10; Length 447;
 Query Match
  Best Local Similarity 77.3%; Pred. No. 5.3e-60;
  Matches 297; Conservative 0; Mismatches 87; Indels
        626 TTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTT 685
QУ
            41 TTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGTGACAAAAAGACAGTCAC 100
\Gamma0
        686 TAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATA 745
Qу
            101 TGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCACAATTATGACAAAATACA 160
Db
```

```
746 TCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACA 805
Qy
           161 TCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAACA 220
Db
        805 TCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAAACACAGC 865
QУ
           221 TCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAATCCTAACAAGACGCAGC 280
Db
        866 CTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCC 925
Qy
               281 CCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGAGTTCCTCAGCAAGTTTC 340
Db
        926 AAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGA 985
QY
                                      341 AGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACCTATTTAGTTAAACAGA 400
Db
        986 TCCGAGACTTGAAGAAAACGGCCC 1009
Qу
            401 TCAGGGATTTGAAGAGACCAGCTC 424
DU
RESULT 3
US-09-867-701-4953
: Sequence 4953, Application US/09867701
 Patent No. US20020132237A1
 GENERAL INFORMATION:
  MPPLICANT: Aglate, Paul A.
  APPLICANT: Jones, Robert
  APPLICANT: Harlocker, Susan L.
  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
  TEXTLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
  FILE REFERENCE: 210121.497
  CURRENT APPLICATION NUMBER: US/09/867,701
  WERRENT FILING DATE: 2001-05-29
  SUMBER OF SEQ ID NOS: 10912
  @OFTWARE: FastSEQ for Windows Version 4.0
  SEQ. ID NO 4953
   LENGTH: 450
   TYPE: ONA
   ORGANISM: Homo sapien
US-09-867-701-4953
                            Score 244.8; DB 10;
                                              Length 450;
                     24.1%;
 Query Match
                     77.3%; Pred. No. 5.3e-60;
 Best Local Similarity
                                                         Gaps .
                           0; Mismatches
                                         87;
                                             Indels
 Matches 297; Conservative
        626 TTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTT 685
QУ
                  27 TTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGTGACAAAAAGACAGTCAC 86
ηb
        636 TAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATA 745
Qу
            37 TGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCACAATTATGACAAAATACA 146
Do
        745 TCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACA 805
Qy.
            147 TCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAACA 206
D'n
```

```
806 TCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGC 865
QУ
          207 TCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAATCCTAACAAGACGCAGC 266
Db
       856 CTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCC 925
Qу
          267 CCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGAGTTCCTCAGCAAGTTTC 326
Db
       926 AAAAAGAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGA 985
Qy ...
          327 AGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACCTATTTAGTTAAACAGA 386
Db
       986 TCCGAGACTTGAAGAAAACGGCCC 1009
Qy
          387 TCAGGGATTTGAAGAGACCAGCTC 410
Db
RESULT 9
US-09-910-943-35
; Sequence 35, Application US/09910943
 Patent No. US20020081610A1
 GENERAL INFORMATION:
  APPLICANT: Hemmati-Brivanlou, Ali
  APPLICANT: Altman, Curtis
  THITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
  FILE REFERENCE: 7529/1G148US1
  CURRENT APPLICATION NUMBER: US/09/910,943
  CURRENT FILING DATE: 2001-07-23
  MUMBER OF SEQ ID NOS: 742
  SOFTWARE: PatentIn version 3.1
 SEQ ID NO 35
   LENGTH: 762
   TYPE: DNA
   ORGANISM: Xenopus laevis
   FEATURE:
   NAME/KEY: misc feature
   LOCATION: (1)..(762)
   OTHER INFORMATION: n may be a or g or c or t/u
U3-09-910-943-35
                          Score 210.8; DB 9; Length 762;
                    20.8५;
  Query Match
                    78.6%; Pred. No. 4.8e-50;
  Best Local Similarity
                                                           3;
                         0; Mismatches 75;
                                          Indels
  Matches 287; Conservative
         1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
Qγ
           397 ATGAAGAAAATGCCATTGTTCAGCAAGTCACATAAAAATCCGGCTGAGATTGTTAAAACT 456
Db
         Çу
           457 CTGAAGGACAACATGGCCCTGCTGGAAAGGCAGGACAAAAAAACTGAAAAGGCCTCTGAA 516
Dio.
        Qy
           517 GAAGTGTCTAAATCTCTTCAAGCTACAAAAGAGATTTTGTGTGGGACAGGGGACAAAGAA 576
DD
```

```
181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
QУ
            577 CCTCAGACAGAGACGGTGGCTCAGCTCGCACAAGAACTGTACAACAGTGGCTTGTTGGTT 636
ďQ
         241 ACACTGATAGCTGACC-TGCAGCTGATAGACTTTGAGGGAAAAAAAAGATGTGACCCAGAT 299
QУ
            637 ACTTTAATAGCCCACCTTGCATCTCATAGATTTTGANGGCAAGAAAGATGTATCTCAGAT 696
Dio
         300 ATTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAG 359
Сλ
            697 ATTCNAC-ACATCCTGAGAAAACAGATTGGCACTCGGAGTNC-CCTGTGGAGTATATCAA 754
Db
         350 TGCTC 364
            1 1 1
         755 TTCCC 759
RESULT 10
US-09-954-456-1453/c
; Sequence 1453, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
  APPLICANT: Young, Paul
   TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents
Using Cancer Gene
FREE OF INVENTION: Sets
   YELE REFERENCE: 689290-76
   CORRENT APPLICATION NUMBER: US/09/954,456
   CURRENT FILING DATE: 2001-09-18
   PRIOR APPLICATION NUMBER: US/60/233,617
   MRIOR FILING DATE: 2000-09-18
   PRIOR APPLICATION NUMBER: 168/60/234,052
   PRIOR FILING DATE: 2000-09-20
   PRIOR APPLICATION NUMBER: US/60/234,923
   PRIOR FILING DATE: 2000-09-25
   PRIOR APPLICATION NUMBER: US/60/235,134
   PRIOR FILING DATE: 2000-09-25
   PRIOR APPLICATION NUMBER: US/60/235,637
   PRIOR FILING DATE: 2000-09-26
   PRIOR APPLICATION NUMBER: US/60/235,638
   PRIOR FILING DATE: 2000-09-26
   PRIOR APPLICATION NUMBER: US/60/235,711
   PRIOR FILING DATE: 2000-09-27
   PRIOR APPLICATION NUMBER: US/60/235,720
   PRIOR FILING DATE: 2000-09-27
  PRIOR APPLICATION NUMBER: US/60/235,840
   PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,863
   PRIOR FILING DATE: 2000-09-27
   NUMBER OF SEQ ID NOS: 2276
   SOFTWARÉ: PatentIn version 3.0
  SEO ID NO 1453
    LENGTH: 387
    TYPE: DNA
    ORGANISM: Homo sapiens
```

US-09-954-456-1453

```
19.2%; Score 195; DB 10; Length 387;
 Ouery Match
                   100.0%; Pred. No. 1.1e-45;
 Best Local Similarity
                                                     Gaps
                                                  0;
                        0; Mismatches
                                       0;
                                          Indels
 Matches 195; Conservative
       820 TTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 879
          ÖУ
       387 TTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 328
Db
       Qγ
           Db
       940 GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 999
           Ωy
       267 GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 208
Db
       1000 AAAACGGCCCCTTGA 1014
Q7
           20% AAAACGGCCCCTTGA 193
D.D
RESULT 11
US-09-550-107-481/c
; Sequence 481, Application US/09880107
: Datent No. US20020142981A1
 CHUERAL INFORMATION:
  CAPPLICANT: Horne, Darci T.
  PARATCANT: Vockley, Joseph G.
  APPLICANT: Scherf, Uwe
  APPLICANT: Gene Logic, Inc.
                           . .
  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
  FILE REFERENCE: 44921-5028-WO 00 20
  CURRENT APPLICATION NUMBER: US/09/880,107
  CURRENT FILING DATE: 2001-06-14
  PRIOR APPLICATION NUMBER: US 60/211,379
  PRIOR FILING DATE: 2000-06-14
   PRIOR APPLICATION NUMBER: US 60/237,054
   PRIOR FILING DATE: 2000-10-02
  NUMBER OF SEQ ID NOS: 3950
   SOFTWARE: PatentIn Ver. 2.1
  SEQ ID NO 481
   EENGTH: 387
   TYPE: DNA
    ORGANISM: Homo sapiens
    FEATURE:
    OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA234362
US-09-330-107-481
                                          Length 387;
                    . 19.2%; Score 195; DB 10;
  Query Match
  Post Local Similarity 100.0%; Pred. No. 1.1e-45;
                                                      Gaps
                          0; Mismatches
                                           Indels
                                        0;
  Matches 195; Conservative
        820 TTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 879
            QУ
        337 TTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 328
 Db
         Qу
```

: E

```
Db
        940 GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 999
Qу
            267 GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 208
Db.
       1000 AAAACGGCCCCTTGA 1014
QУ
            1111111111111
        207 AAAACGGCCCCTTGA 193
Db
RESULT 12
US-10-257-826A-113
; Sequence 113, Application US/10257826A
 Publication No. US20030181407A1
 GENERAL INFORMATION:
  APPLICANT: SA MAJESTE LA REINE DU CHEF DU CANADA
  APPLICANT: PALIN, Marie-France
  APPLICANT: POMAR, Candido
  APPLICANT: GARIEPY, Claude
  TITLE OF INVENTION: Steatosis-modulating factors and uses
  TITLE OF INVENTION: thereof
  FILE REFERENCE: 14654-2US
   CURRENT APPLICATION NUMBER: US/10/257,826A
  CURRENT FILING DATE: 2002-10-17
  PRIOR APPLICATION NUMBER: 60/197936
   PRIOR FILING DATE: 2000-04-17
   WRIGH APPLICATION NUMBER: PCT/CA01/00509
  SPRICE FILING DATE: 2001-04-12
                                                            1365
   MARKER OF SEQ ID NOS: 305
                                                            1500
   SOFTWARE: FastSEQ for Windows Version 4.0
  (200 ID NO 118
    NGTH: 722
    YPE: DNA
    CRGANISM: Artificial Sequence
    OTHER INFORMATION: Artificial sequence
    OTHER EMFORMATION: Muscular steatosis
    OTHER INFORMATION: Porcine
    FEATURE:
    NAME/KEY: misc feature
    LOCATION: (1)...(722)
    OTHER INFORMATION: n = A, T, C or G
US-10-257-826A-118
                       16.7%; Score 169.8; DB 13;
                                                 Length 722;
  Query Match
                       60.1%; Pred. No. 3.1e-38;
  Best Local Similarity
                                                Indels
                             0; Mismatches 196;
  Matches 303; Conservative
         347 TGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCC 406
 Qу
                        2 TGGTGAATNCCTCTGCCCCCACNGAATTTTTGGTCATGGTANTNGAAGGGGATNAATNTT 67
 Db
         407 CACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTG 466
 \Omega Y
             68 CCGAAATTNCNNTTAATTGGGGNATNATGGTNAGANAATGCCTTNGACCTNNACCGCTTG 127
 Db
```

```
467 CCAAAATCATCCTCTTTTC--TAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTC 524
QУ
           128 CCAAAATCATTTTGNGGGCCGAACACAGTTTATAGAGATCTTCACATATGTCTAAATGTN 187
Db
        525 AACA-TTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATA 583
CA
           138 ANCATTTTNACATATCTTTACATNCNNTTNCCNCATTTTNNGNNTTACTTTCACGACATA 247
ďŒ
        584 AAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGA 643
Ov
                   248 TATTGCTCACNGCGCAANTTTTGGAACANCATTATGATANATTTTTCAGTGAATATGATG 307
Db
        644 AATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGC 703
QY
           308 AAGNGCTTCATTCTTAAAATTATGTGGCCACAAGACAATCACTGAAGCTTCTCGGNGAAC 367
Db
        704 TGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACC 763
Qν
           368 TACTACTANATAGACNCNACTTCNCCANTATGACCACATACCTCATTAAACCTGNGNACC 427
Db
        764 TGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAA-CATCCAGTTTGAAGCCTTT 822
Qy.
            438 T-CCATTAATGATGAACCTGCCTGCAGACAAAAGTCGGAACCTTCCANTTTGAGGGCTTN 486.
D)
       . 823 CATGTTTTAAGGTGTTTGTGGCC 846
Qt:
            487 CACGTTTTTAANGGGGNTGTNNNC 510
AMSKEIT 13
UB-10-257-326A-119
; Sequence 119, Application US/10257826A
: Publication No. US20030181407A1
; GENERAL INFORMATION:
 APPLICANT: SA MAJESTE LA REINE DU CHEF DU CANADA
  APPLICANT: PALIN, Marie-France
   APPLICANT: POMAR, Candido
   APPLICANT: GARIEPY, Claude
   TITLE OF INVENTION: Steatosis-modulating factors and uses
   TITLE OF INVENTION: thereof
   FILE REFERENCE: 14654-2US
   CURRENT APPLICATION NUMBER: US/10/257,326A
   CURRENT FILING DATE: 2002-10-17
  FRIOR APPLICATION NUMBER: 60/197936
   PRIOR FILING DATE: 2000-04-17
   PRIOR APPLICATION NUMBER: PCT/CA01/00509
   PRIOR FILING DATE: 2001-04-12
   NUMBER OF SEQ ID NOS: 305
   SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 119
    LENGTH: 700
    TYPE: DNA
    ORGANISM: Artificial Sequence
```

FEATURE:

OTHER INFORMATION: Artificial sequence OTHER INFORMATION: Muscular steatosis

```
OTHER INFORMATION: Porcine
   FEATURE:
   NAME/KEY: misc_feature
   LOCATION: (1)...(700)
   OTHER INFORMATION: n = A, T, C or G
US-10-257-826A-119
                    16.4%; Score 166.6; DB 13; Length 700;
 Query Match
                    60.1%; Pred. No. 2.6e-37;
 Best Local Similarity
                          0; Mismatches 197; Indels
                                                    5; Gaps
 Matches 304; Conservative
        345 TGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGC 404
Qy
                            1111
         5 TCTGGTGAATCCCTCTGCCCCCACNGAATTTTTGGTCATGGTANTNGAAGGGGATNAATN 65
Db
        405 CCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACT 464
             QV
         66 TTCCGAAATTTCGATTAATTGGGGNATNATGGTNAGANAATGCCTTNGACCTCCACCGCT 125
Db
        465 TGCCAAAATCATCCTCTTTTC--TAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTG 522
Qγ.
                         126 TGCCAAAATCATTTTGNGGGCCGAACACAGTTTATAGAGATCTTCACATATGTCTAAATG 185.
Db
        TCAACA-TTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACA 581
Qy
           306 S'NANCATTTTNACATATCTTTACATNCNNTTNCCNCATTTTNNGNNTTACTTTCACGACA 245
        HAVE TAAAGTGTTGGTAGCAGACTTCTTAGAACAAATTACGACACTATTTTTGAAGACTATGA 6418
24.6 TATATTGCTCACNGCGCAANTTTTGGAACANCATTATGATANATTTTTCAGTGAATATGA 305.4
1.15
        42 GAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGA 701
             396 IGAAGNGCTTCATTCTTAAAATTATGTGGCCACAAGACAATCACTGAAGCTTCTCGGNGA 365
图为
        702 GCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAA 761
QУ
            356 ACTACTACTANATAGACNCNACTTCNCCANTATGACCACATACCTCATTAAACCTGNGNA 425
700
        762 CCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAA-CATCCAGTTTGAAGCCT 820
                Сy
        426 CCT-CCATTAATGATGAACCTGCCTGCAGACAAAAGTCGGAACCTTCCANTTTGAGGGCT 484
Db
         821 TTCATCTTTTTAAGGTGTTTTGTGGCC 846
 ÖУ
            485 TNCACGTTTTTAANGGGGNTGTNNNC 510
Db
 RESULT 14
 US-09-770-445-592
 ; Sequence 592, Application US/09770445
  Patent No. US20020023281A1
  GENERAL INFORMATION:
   APPLICANT: Gorlach, Jorn
   APPLICANT: An, Yong-Qiang
   APPLICANT: Hamilton, Carol M.
```

APPLICANT: Price, Jennifer L.

```
APPLICANT: Raines, Tracy M.
           Yu, Yang
 APPLICANT:
            Rameaka, Joshua G.
 APPLICANT:
            Page, Amy
  APPLICANT:
            Matthew, Abraham V.
  APPLICANT:
            Ledford, Brooke L.
  APPLICANT:
            Woessner, Jeffrey P.
  APPLICANT:
  APPLICANT: Haas, William David
  APPLICANT: Garcia, Carlos A.
  APPLICANT: Kricker, Maja
            Slader, Ted
  APPLICANT:
            Davis, Keith R.
  APPLICANT:
            Allen, Keith
  APPLICANT:
            Hoffman, Neil
  APPLICANT:
            Hurban, Patrick
  APPLICANT:
  TITLE OF INVENTION: Expressed Sequences of Arabidopsis
  TITLE OF INVENTION: thaliana
  FILE REFERENCE: 2023US (PARA-012PRV)
  CURRENT APPLICATION NUMBER: US/09/770,445
  CURRENT FILING DATE: 2001-01-26
  PRIOR APPLICATION NUMBER: US 60/178,472
  PRIOR FILING DATE: 2000-01-27
  NUMBER OF SEQ ID NOS: 999
  SOFTWARE: FastSEQ for Windows Version 4.0
 CEQ ID NO 592
   LENGTH: 861 ·
   TYPE: DNA
  OEGANISM: Arabidopsis thaliana
ya-69-770-445-592
                      15.4%; Score 156; DB 9; Length 861;
 Query Match,
                      55.8%; Pred. No. 3.4e-34;
 Test Local Similarity
                           0; Mismatches 250; Indels
                                                          Gaps
 Matches 319; Conservative
        394 GGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGA 453
            12 GGGTTTGAAAACACCGATATGGCGTTACACTATGGTACTATGTTTAGAGAGTGCATCCGT 71
Db
        454 CATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTAC 513
Q_{\underline{J}'}
                    72 CATCAGATTGTTGCAAAATATGTTTTGGACTCGGAGCACGTGAAGAAGTTTTTTTACTAC 131
Db
        514 GTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTA 573
97
             132 ATACAGCTTCCCAATTTCGACATTGCTGCTGATGCTGCTGCAACTTTTAAGGAACTTCTG 191
Tito
        33 ACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAA
25
                        11 11 11 11
        192 ACAAGGCACAAGTCTACAGTTGCTGAGTTTCTCATTAAGAATGAAGACTGGTTTTTTGCA 251
D'5
        634 GACTATGA---GAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAG 690
QУ
                                                252 GACTACAACTCAAAGCTTCTTGAATCAACTAATTATTACCCGACGGCAAGCTATTAAG 311
Db
        691 CTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGC 750
QУ
             312 TTGTTGGGCGATATATTATTGGATAGGTCAAATTCGGCTGTGATGACGAAGTATGTGAGC 371
Db
```

```
751 AAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAG 810
             Qy
        Db
        811 TTTGAAGCCTTTCATGTTTTTAAGGTGTTTTGTGGCCAGTCCTCACAAAACACAGCCTATT 870
            QY
        432 ATAGAAGCTTTCCATGTTTTCAAGCTGTTTGTAGCGAACCAAAACAAGCCTTCAGACATC 491
Db
        871 GTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAA 930
Qу
                             492 GCCAACATTCTGGTGGCAAACAGAAACAAGCTTCTGAGATTGTTGGCTGATATCAAGCCG 551
Db
        931 GAAAGGACGGATGATGAGCAGTICGCTGACGA 962
QУ
            552 GACAAAGAGGACGAGAGGTTTGACGCAGACAA 583
Db
RESULT 15
US-09-923-876-1251
; Sequence 1251, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
  MPPLICANT: Lalgudi, Raghunath V.
  APPLICANT: Kamigaki, Laura Y. (Ito)
  APPLICANT: Sherman, Bradley K.
  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN
                                                        双翼 (Trans) (1)
ATTIMINING
                                                          43.33.1.2
  COMMERCE REFERENCE: PL-0012-1 CON
 MADERENT APPLICATION NUMBER: US/09/923,876
  WERRENT FILING DATE: 2001-08-06
  FINOR AFPLICATION NUMBER: 09/298,329
   CRIOR FILING DATE: 1999-04-21
   FRIOR APPLICATION NUMBER: 60/085,331
  PRIOR FILING DATE: 1998-05-05
  NUMBER OF SEQ ID NOS: 6332
  SOFTWARE: PERL Program
  SEC ID NO 1251
   LENGTH: 262
    TYPE: DNA
   ORCANISM: Zea mays
    FEATURE:
   MAME/KEY: misc feature
   OTHER INFORMATION: Incyte ID No. US20020013958A1 700158378H1
    MAME/KEY: unsure
  LOCATION: 148
 , CTHEE INFORMATION: a, t, c, g, or other
 US-09~923-876-1251
                       7.3%; Score 74.2; DB 9; Length 262;
  Quary Match
  Best Local Similarity 55.5%; Pred. No. 6.4e-11;
  Matches 142; Conservative 0; Mismatches 114; Indels
         311 TCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTC 370
 Qv
             7 TGTTGAGACAGAAGGTTGATGAAAGCTATTGTTGCGTCCAGTATATTGAAAATCATTTTG 66
 D'o.
```

QУ	371 ATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGA 430	
Db	67 ATCTTTTGGATTTCCTTGTTTGCTATAAGAACTTGGAAGTCGCGTTGAATTGTGGAA 126	
QУ	431 TTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATC 490	
Db	127 ACATGTTGCGAGAATGCATAANATATCCTACACTTGCAAAATATATTTGGAGTCAAGCA 186	
ΩΥ	49% AATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCT 550	
Dр	187 GCTTCGAGTTGTTTTCCAGTATGTTGAATTGTCAAACTTCGATATTGCATCTGATGCTC 246	
QУ	551 TTGCTACTTTCAAGGA 566	
Бb	247 TGAACACTTTCAAGGA 262	•

Search completed: January 6, 2004, 05:04:45
Job time: 1400 secs